Editorial

Congress Chairmen's Preface

The 5th World Congress on Alternatives and Animal Use in the Life Sciences was a memorable event at which 970 colleagues from 46 countries met on August 21-25, 2005, in Berlin (Germany) to discuss progress in refining, reducing and replacing testing in experimental animals according to the Three Rs Principle of Bill Russel and Rex Burch. The Alternative Congress Trust and the organisers and sponsors were happy to welcome delegations of more than 100 participants from Germany, Japan, the UK and the USA and scientists from countries, who were able to attend for the first time, e.g. Cuba, India, Iran and Eastern Europe, most of whom were supported by funding generously provided by the sponsors.

We are convinced that all of the participants enjoyed the relaxed atmosphere and the stimulating discussions at the Estrel Convention Centre, and also the pleasant summer evenings at different locations in Berlin, the unified and re-established German capital, which provided a perfect setting to refresh memories with old friends and to meet young colleagues from around the world. We were fortunate that two charismatic women accepted our invitation to give the opening speeches, Jane Goodall and Renate Künast. Both of them fascinated the international audience with the power of their arguments that needed no advanced technical support. During Jane Goodall's lecture, everyone in the audience, from students to senior scientists and managers, hung on her lips to catch every word. Renate Künast, the Patron of the congress and German Federal Minister for Consumer Protection, Food and Agriculture, stressed the firm commitment of the German government to animal welfare, including experimental animals, and to the Three Rs principle. These were the most memorable moments of the congress. The two Europeans encouraged scientists from all continents to follow their example. Participants representing countries that cannot yet give the Three Rs a high priority due to economic constraints expressed their appreciation to Jane Goodall and Renate Künast for setting high and realistic ethical standards in the life sciences.

The European Union gives the Three Rs concept a very high priority, as demonstrated e.g. in the EU cosmetics directive and in the new EU chemicals policy (REACH). Therefore, the European cosmetics and chemicals industry, international regulatory agencies and animal welfare organisations chose the 5th World Congress in Berlin as a scientific forum to discuss progress in safety testing without animals. Scientists from academia and industry presented advanced molecular and cell culture methods that will in the near future even allow replacement of experiments in transgenic animals.

The Scientific Committee had selected the following 7 themes, "education", "laboratory animal welfare", "moral issues of animals, alternatives and public policy", "information systems and data bases", "safety testing, validation and risk assessment", "mathematical modelling" and "applying new science and technology". In addition to the invited plenary lectures more than 600 abstracts were discussed in workshops and as oral or poster presentations. It is remarkable that 10% of the abstracts were submitted to the session of *in vitro* testing for skin irritation. This shows

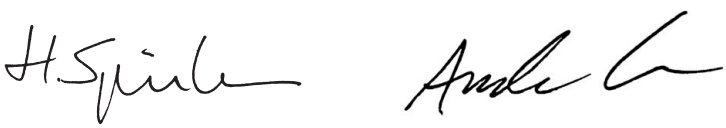
the high priority that scientists in academia and industry give the EU Cosmetics Directive, which calls for a ban of safety testing of cosmetic ingredients in animals in the very near future. We are also pleased that the European Commission, the European chemical and cosmetics industry and international animal welfare organisations used the 5th World Congress in Berlin as their forum to discuss ethical, scientific and political aspects of reducing safety testing in animals according to the Three Rs principle.

The organisers highly appreciate the tradition that the respected international institutions used the 5th World Congress as a venue to award prizes devoted to scientific and ethical aspects of the Three Rs. In fact, more than 15 prizes were awarded at the plenary sessions and during social events of the 5th World Congress in Berlin.

We would not have been able to offer such an attractive program without the substantial support of international sponsors. We were honoured that the congress was held under the patronage of the German Minister of Consumer Protection, Food and Agriculture Renate Künast and we are proud that the institutions sponsoring the 5th World Congress include the Federal US Government (NIH) and the UK Home Office as well as major global companies from the pharmaceuticals, chemicals and cosmetics industry and international animal welfare organisations.

We are also indebted to the congress organiser CTW, in particular Thomas Wiese, Martina Falliner-Dieterici and Wencke Wieseke, who supported the local organisers in Berlin during the period of planning and behind the scenes during the conference sessions and the social events in the evening, among which the gala dinner at the greenhouse in the botanical gardens was the most memorable. We are also indebted to the BfR, the Federal Institute for Risk Assessment, where the staff of ZEBET, and in particular the scientific secretary Richard Vogel, worked very hard to provide an attractive setting for the 5th World Congress.

Finally, we appreciate the co-operation with Thomas Hartung, head of ECVAM, the European Centre for the Validation of Alternative Methods in Ispra (Italy), which provided funding to publish the Proceedings of the 5th World Congress on Alternatives and Animal Use in the Life Sciences, and we thank Franz Gruber and the editorial office of ALTEX in Zurich (Switzerland), who did an excellent job and worked long hours to provide you with these proceedings.



Horst Spielmann and Andrew Rowan
Co-Chairs of the Organising Committee

Proceedings of the 5th World Congress on Alternatives, Berlin 2005

Contents

Theme 4 Information systems and databases

4.1 Session: 3Rs database and services – developments worldwide

Annett J. Roi and Manuela Flego: ECVAM's database service on alternative methods (Db-ALM) – online

4.2 Session: Information retrieval – search strategies and search engines

Hua Florence Chang: Alternative search methods to retrieve information on the web

Barbara Grune, Amrei Schnock, Antje Dörendahl, Susanne Skolik and Horst Spielmann: Searching strategies for detecting publications on alternative methods: A pilot study

Jean A. Larson: The animal welfare information center: Helping the community meet the information requirement of the 3Rs

Mary W. Wood and Lynette A. Hart: Considering animal alternatives and welfare via a comprehensive search of the scientific literature

4.3 Workshop: Search strategies – user requirements

Jodie A. Kulpa-Eddy: Overview of the regulatory requirements for the consideration of alternatives

Theme 5 Safety testing, validation and risk assess

5.1 Session: Strategies for using non-animal methods in relation to chemicals legislation (HPV, REACH) (ECVAM-Session)

Wolf-Rüdiger Bias: REACH and CEFIC's conception of a feasible, information and priority based approach

Uwe Lahl: Strategy for minimising the use of animal testing as part of REACH

Marilyn Wind, Bradley Blackard, Raymond Tice, William Stokes, David Hattan and Leonard Schechtman: ICCVAM's role in validating *in vitro* test methods for endocrine disruptor screening*

5.2 Session: New approaches to risk assessment (ESTIV-Session)

Michael Gülden and Hasso Seibert: In vitro-in vivo extrapolation of toxic potencies for hazard and risk assessment – problems and new developments

5.3 Session: Progress and needs for developing and validating alternatives for dermal sensitisation testing

Pierre Aeby, David Basketter, Walter Diembeck, Frank Gerberick, Hiroshi Itagaki, Ian Kimber, Béatrice Le Varlet, Irene Manou, Marc Paye, Françoise Rousset, Joanna Rowland and Hitoshi Sakaguchi: Colipa dendritic cell research projects

Maja Divkovic: Hapten-protein binding: What do we know?

Jean-Pierre Lepoittevin: The chemistry of skin allergy

5.5 Session: Advancements and needs for developing and validating 3R alternatives for ocular irritancy testing

Matthias Herzler, Horst Spielmann, Ingrid Gerner, Manfred Liebsch and Thomas Hoefer: Use of *in vitro* data and (Q)SARs to classify eye irritating chemicals in the EU – experience at the BfR

Jill Merrill, Karen Hamernik, Leonard Schechtman, William Stokes and Marilyn Wind: ICCVAM progress in evaluating *in vitro* test methods for identifying severe ocular irritants/corrosives

*This paper reflects the views of the authors, has not been reviewed or approved by, and may not necessarily reflect the views of the Consumer Product Safety Commission or other agencies. This work was supported in part by NIEHS Contract N01-ES 35504. Since this paper was written as part of the official duties of the authors, it can be freely copied.

*Summarised from the poster *Animal carcinogenicity studies: poor human predictivity* by Andrew Knight, Jarrod Bailey and Jonathan Balcombe, which received the Animal Welfare Poster Award from Deutscher Tierschutzbund (the German Animal Welfare Federation) at the 5th World Congress on Alternatives and Animal Use in the Life Sciences, Berlin, 25 August 2005. Reproduced with permission from the complete paper by Knight and Bailey (2005) in *Alternatives to Laboratory Animals*.

Opening Speeches

Opening address

Renate Kuenast
(at the time of the 5th WC German Federal Minister of Consumer Protection, Food and Agriculture)

Ms. Goodall,
Mr. Spielmann,
Ladies and gentlemen,

Welcome to the 5th World Congress on Alternatives and Animal Use in the Life Sciences! I am proud to have the opportunity to follow in the footsteps of Baltimore, Utrecht, Bologna and New Orleans and carry on this new and important tradition in animal welfare here, today, in Berlin!

Animal welfare is very important to us in Germany – in all areas:
- for livestock on farms and animals in the wild,
- for animals during transport and slaughter,
- for pets as well as laboratory animals.

We have established a sound legal foundation for this by incorporating animal welfare as a state aim into our constitution. We succeeded in amending the German constitution to include animal welfare as a state aim only because broad segments of German society are interested in animal welfare issues.

And because there is a steadily growing sense of responsibility for animals as fellow creatures among our citizens. This sense of responsibility has grown over the years. Consumer interest in animal welfare wasn't always as strong as it is today – not all animal welfare issues have been the focus of public discussion to the same degree. In fact, one issue has been anchored in the public's awareness for decades: Namely, the issue of animal testing! The credit for this goes first and foremost to the work of numerous organisations.

For more than 30 years now, animal welfare organisations have focused their criticism on animal testing. And scientists, industry and special interest organisations have been getting together every three years now for the last 12 years to discuss recent advances and new developments.

Ladies and gentlemen:
The 5th World Congress for Alternatives and Animal Use in the Life Sciences has developed into an international forum on animal welfare!

We all share the same goal: To reduce the suffering of animals as much as possible! The best and most desirable solution – following Russel and Burch's three "R" concept – is the replacement method: in other words, the best and most desirable solution would be to replace animal testing completely. But we also know that there are areas where we cannot completely dispense with animal testing. However – in keeping with the spirit of reduction – we want to at least substantially reduce the number of animals being used. Furthermore, we want to reduce the suffering of laboratory animals and the harm done to them as much as possible.

Ladies and gentlemen:
Animal welfare is an important intrinsic value. The challenge and obligation that we face here is to always look very closely when we weigh conflicting interests. We must ask ourselves: Where is animal testing truly indispensable for making our lives safer? Look at the pharmaceutical industry for example. And on the other hand: How can we do the most to replace animal testing? The underlying issue here is the value of life. This is an ethical question to which there are definitely no simple answers. One thing however must be understood: We cannot allow these two values – consumer protection and animal welfare – to be played off against one another! Instead, our goal must be to rec-

oncile them. I know that this is precisely the standard that you all set for yourselves and your work – and I would like to take this opportunity to thank you for this!

Ladies and gentlemen:
Animal welfare needs you. Although much progress has been made in recent years to help animals, we have to admit that animal welfare is still far from being a matter of course! This became clear, for instance, during the discussion over the draft constitution for the European Union: Establishing animal welfare as a cross-sectoral task has been a major victory for animal welfare. It is also a prerequisite for putting animal welfare right at the heart of the European Union. Looking at the long term however, our objective must be to establish animal welfare as a separate EU goal.

A Europe-wide survey conducted in June showed that the time is ripe for this. A majority of the persons surveyed in 15 of the EU's 25 member states said that they would like to see effective animal welfare in their country. A total of 46 percent of European consumers said that they would spend between five and ten percent more for products that have been produced in a way that satisfies animal welfare requirements. Eleven percent said they would be willing to spend even more.

We must of course be a little careful when interpreting these figures because what people say is not always the same as what they do. Actually, what is more interesting here is the fact that young people in particular feel that animal welfare is important. It always takes longer for behaviour to change than for awareness to develop.

Consequently, I am convinced that purchasing decisions will start catching up in the near future. The first generation of the 21st century will base its purchasing decisions on a different notion of quality. A notion that is based not only on product quality but also on process quality. In other words: the specific question of how and where products were produced, whether fair wages were paid, whether environmental criteria played a role in their production and – of course – whether due regard was paid to animal welfare. In my opinion, these will all be business advantages in the future.

And, as I would like to emphasise here: These sustainability-related concerns will become increasingly important for trade as consumer awareness of the importance of animal welfare grows. This is why it doesn't make sense to me that the World Trade Organization still classifies these criteria as "non-trade" concerns. The WTO will have to change its stance on this. The German government will continue its efforts to get animal welfare finally recognised.

In light of recent events, I would like to emphasise how important animal welfare is just one more time. We live in a world that is marked by ever-growing globalisation. As a result, animal health has taken on an entirely new meaning for everyone involved. The outbreak of avian flu in Asia has made it all too clear that epidemics among animals do have a major impact on trade. We live in an age of growing global flows of goods. This makes international solutions a necessity. I consider it a big step forward that the OIE has recognised animal welfare as an essential factor in fighting such epidemics. Now we need to take that second step – namely, we need to establish animal welfare as a trade concern at the WTO.

Ladies and gentlemen:
The fact that trade and animal welfare belong together is also evidenced by trade and industry's involvement in this conference. The cosmetic industry is setting a good example here. In fact, in Germany, animal welfare was incorporated into the cosmetics industry years ago. As a rule, Germany does not allow animal tests to be used in the development of tobacco products or cosmetics. The EU Cosmetics Regulation has further improved the protection of animals in the testing of such products.

The Directive was transposed into national law in Germany in 2004. The new legislation prohibits the sale of cosmetic products that were tested on animals to ensure compliance with the provisions of this Directive – despite the existence of alternative methods that have been recognised by the European Union and the OECD. Starting on January 1st 2009, Community law will prohibit the sale of cosmetic products containing ingredients that were tested on animals after this date in order to meet the safety requirements of cosmetics regulations. This will apply even when there are no alternative testing methods. An exception has been made in four specific cases and the deadline for these was pushed back to March 2013. Given this situation, it is gratifying that the cosmetics industry is using the 5th World Congress here in Berlin as a forum: presenting to the scientific community recent advances in the development of new methods for testing the toxicological safety of cosmetic products.

colipa – the European Cosmetic, Toiletry and Perfumery Association – is intensively involved in the 5th World Congress. In addition, leading manufacturers of cosmetic products and their ingredients are important sponsors for this conference. A number of firms will also be presenting research projects that they have conducted to develop non-animal testing methods. It comes as no surprise that more than 10 percent of the abstracts submitted for lectures deal with in vitro methods for testing skin tolerance.

Ladies and gentlemen:
We also have to talk about the fact that although Germany has banned the use of animal tests in the development of cosmetics and tobacco products, such tests are required by law in other areas. Chemicals legislation is the source of one highly topical issue. The European Commission submitted its proposal for a regulation for the introduction of an EU-wide System for the Registration, Evaluation and Authorisation of Chemicals (also known as the REACH system) in 2003. This proposal contains the following key points:

- Manufacturers must register any substance that they produce in quantities of one or more tons a year. This requirement would apply to some 30,000 substances.
- A single system will apply to existing and new substances and the data requirements for both will be based on the existing procedure for new substances.
- Industry will be assigned greater responsibility, with government agencies concentrating on substances that are of very high concern and/or are produced in large quantities.

- It will be possible to require an authorisation procedure for some substances of very high concern.
- Downstream users will be incorporated into the system in cases where they do not wish to disclose their use of a particular substance to the manufacturer. The pending Community provisions will require existing substances to be comprehensively tested as well. This will improve our data and strengthen health, consumer and environmental protection interests considerably. And this is precisely what makes weighing these interests such a difficult balancing act.

I personally feel that an increase in the number of laboratory animals that is limited to the registration period for existing substances is justified when absolutely all scientifically satisfactory means for avoiding unnecessary animal tests have been exhausted. Looking at the Commission's proposal, there is still need for improvement in this area. Possible improvements include:

- Adding other already validated alternative methods to the Regulation's annex,
- Supporting the development and validation of alternative methods and
- Avoiding unnecessary duplication of testing activities, particularly by including a provision on using test documentation, following the provisions of the German Chemicals Act. This will require everyone to pull together:
- The scientific community will have to work together on the development of alternative methods and their validation.
- However, the political sector will also have to support the rapid development of alternatives to animal tests.

Ladies and gentlemen:
I would like to briefly outline the situation in Germany for you: The German government has set up two funding programs:

- The first is the Methods to Replace Animal Testing funding priority at the Federal Ministry of Education and Research that has been in place since 1984.
- The second is the Allocation of Research Funding for the Scientific Development of Alternatives to Animal Experiments programme which ZEBET – the Centre for Documentation and Evaluation of Alternatives to Animal Experiments at the Institute for Risk Assessment – has managed since 1990.

I would like to take this opportunity to expressly thank Professor Spielmann, the head of ZEBET, for his dedicated work. Both programmes are aimed at developing alternative methods to the toxicological safety testing on animals that the law requires. ZEBET frequently funds the groundwork done at individual laboratories to develop new methods. When this work shows promise, the Federal Ministry of Education and Research then funds pre-validation studies or validation studies in cooperation with industrial partners. Through these funding programmes, the Federal Republic of Germany wants to make what will be the largest contribution by far among all EU member states toward the development of non-animal testing methods.

And last of all, we have a third instrument to mirror the high scientific standards placed on the development of alternatives to animal testing. The Animal Welfare Research Prize, which is awarded every year together with prize money of 15,000 euros, seeks to encourage outstanding research work that revolves around methodological studies aimed at limiting and replacing animal testing. I am pleased to have the opportunity to single out this outstanding congress with this award later on.

I wish you all a very productive conference and a pleasant stay here in my hometown of Berlin!

Correspondence to
Renate Künast
Member of Parliament
Platz der Republik
11011 Berlin
Germany
www.renate-kuenast.de
renate.kuenast@bundestag.de

Ending Research on Non-Human Primates

Jane Goodall
Introduced by Andrew N. Rowan

Well, first of all, good morning everybody. I understand there are people here from 46 countries, is that right? That means there are 46 languages, or at least nearly 46 languages, represented. But, I want to bring a voice into this room, a different language, I want to introduce a voice of one of those animals whose future we are discussing. And there are people here in this audience who can reply, so that we can swell the sound of the chimpanzees' good morning greeting throughout the room, so let's try! Calls like a chimpanzee – Applause.

Before I talk about my vision for the future of chimpanzees and other non-human primates – and indeed all primates including ourselves, because we are part of the primate family and our future is hanging very much in the balance at the moment as a consequence of our destruction of the environment – I'd like to go back a bit, because Andrew (Rowan) just commented on his feeling of joy that we have come so far, that today we have a minister standing up and speaking eloquently and passionately about animal welfare. So, I'd like to take you back to how things were when I began to study chimpanzees in the very early sixties. When I first went to study the chimpanzees in Gombe National Park in Tanzania, I had no degree of any kind. I was selected by the late Louis Leakey, because he thought he saw in me someone who really cared about animals, who wanted to learn about them, who didn't care about clothes and hairdressing and things like that. So he offered me this opportunity and eventually got the money, and I began. And then, after I'd been in the field about a year, I got this letter from him saying, Jane, I shan't always be here to get money for you. You have to stand on your own two feet. You have to have a degree. We don't have time to mess with a B.A. You'll have to go straight for a Ph.D. I've got you a place in Cambridge University. He said I was going to be studying ethology at Cambridge. I didn't even know what the word meant. There were no faxes and e-mails in those days, so I had to wait to find out. When I got to Cambridge I was excited. I'd just begun to learn about the complexity of chimpanzee society. I'd just begun to come to know some of the individuals and their unique personalities. I was excited to learn about this scientific method. So you can imagine my dismay when I got to Cambridge and was told I had done everything wrong. First of all, I should not have named the chimpanzees. That was very unscientific; they should have had numbers. I couldn't talk about their mind or any possibility of rational thought, because that was unique to us. This was still taught to my son when he was at school, 20 or 30 years ago. And, thirdly, I absolutely could not talk about emotions in any other than human, because they were unique to us, and even if animals other than us did have emotions, we could never prove it. Therefore, we certainly shouldn't talk about it. And so, I was guilty of the terrible sin of anthropomorphism. I also talked about the vivid personalities of the chimpanzees I'd come to know. And that was wrong too. So here I was, young, naïve, no university training, I should have been slightly overwhelmed and subdued and very anxious to do things right.

But, you know, all through my childhood I'd had the most wonderful teacher. A teacher who taught me so much about animals and their behaviour: that they do have personalities and minds and feelings. And that was my dog Rusty. I knew from Rusty that these professors, erudite though they might be, were wrong in this particular instance. And I was very fortunate in having a superb mentor at Cambridge, my thesis supervisor, Robert Hind, who helped me negotiate the tricky scientific pathway of expressing what I believed to be true in such a way that I would not be torn to shreds by the far more scientifically qualified people I was working with. For example, I said to him one day that the old female, Flo, had a baby and that Fifi, the elder sister, was fascinated by this baby. Every time another young chimp came anywhere near, Fifi would chase that youngster away, with all her hair bristling. I said, she was jealous, and Robert said, you can't say that, because you can't prove it. So I said, well she was, so what shall I say? He said, you should say that Fifi behaved in such a way that, had she been a human child, we would say, she was jealous. Now this is brilliant and I took that advice with me throughout my whole career!

To come back to Andrews theme of feeling overjoyed at how far we've come, there are very few major universities in the world today where young people can't actually study the animal mind, animal mentation. We can even study animal emotions and try to sort them out – as well as our own emotions, by the way –, and we can even study animal personality. So, we have come quite a long way, but we haven't yet come far enough.

But, just for a few moments, let me talk about some of the things that have been so fascinating in the 45 years that we've been learning about these amazing chimpanzees at Gombe. Because, in a way, the chimpanzees are so like us that they serve as ambassadors for the rest of the non-human animal species. They show us very clearly that there is no sharp line between us and the rest of the animal kingdom. Back in the early sixties, there were many scientists, many ethologists, who truly believed that there was a line and that there was a disconnect in the process of evolution. There were all the animals, then a line, and then us with our unique qualities of mind and reason, emotions, personality and so forth.

Well, over these years we have learned a great deal about chimpanzees, not only at Gombe, but in other parts of Africa; not only in the wild, but also in some captive situations. And we're beginning to put together a rather awesome picture of this

closest of our known living relatives. About their complex social structure, long term family bonds which can persist through a life of anything up to 70 years, although we'd don't think they live much beyond 50 to 60 in the wild. There is a long childhood, a dependency on the mother during which time the brain is plastic and serves the same useful purposes it does in our own species for learning. We find that as we come up in the animal kingdom to more and more complex brains, so learning plays and ever more important role in acquiring individual behaviour. We find that the chimpanzees show sophisticated cooperation. When we come to non-verbal communication, kissing, embracing, holding hands and – by the way, when I first mentioned kissing and embracing in those early sixties, that was absolutely impermissible! I could not talk about other than humans kissing or embracing, but I did anyway. They kiss, they embrace, they hold hands, they pat one another on the back, and they beg, they shake their fists, they throw rocks. They do these things that we do in the same kind of context, and they clearly mean the same kind of thing. They show tool using behaviour, even tool making behaviour, and I don't know if any of you remember, but at the time in the early sixties, when I first saw a chimpanzee using a tool, I mean, picking a twig and stripping the leaves and making a tool, this had a huge impact on the scientific community, because it was thought that humans, and only humans, used and made tools. That was meant to differentiate us more than anything else from the rest of the animal kingdom. And we have now found out from these other chimp studies across Africa that everywhere where people have watched them in the long term, they have seen different examples of tools and different objects used for different purposes in different ways. And so, because it is very clear that these skills are passed from one generation to the next through observation, learning and imitation, and that chimpanzee youngsters can learn through direct imitation, just like our children, this is a definition of culture. So that we can say that chimpanzees have their primitive culture.

Unfortunately, there is a dark side to their nature. They are capable of extreme brutality, particularly when it comes to interactions between one community and a neighbouring social group. And, sadly, this makes them seem even more human-like than I thought they were before. I was criticised quite heavily when I first published the events I described as a "four year war" to be like some kind of primitive warfare, where the males of one community systematically annihilated the individuals of the neighbouring smaller group. People said that, well, if we think there is a common ancestor then everyone is going to say that aggressive tendencies have been inherited by humans through this ancient primate inheritance, and war and violence are inevitable in our species. Well, I do think we have inherited aggressive tendencies and I dare say that most of the people in this room would agree with me.

But, we have to remember that we've also inherited tendencies of compassion, love and altruism, and these are very clearly demonstrated in the chimpanzees too. And with our incredibly sophisticated intellect, the one thing I feel makes us more unique than any other single thing from the other animals is that we have this extraordinary spoken language, this sophisticated spoken language. This language that we are using today, a language with which we can paint pictures with words and vividly describe things to people who have never been to the place or seen the things we are talking about. We can teach our children about things they don't know. So here we are with our sophisticated intellect, our ability to communicate in ever more sophisticated ways, electronically now, and we are standing at a crossroads as regards our stewardship of this planet, whether it be the environment or whether it be our fellow animals. We really are at a crossroads. And what we are here to talk about today, the three Rs, the eventual elimination of all animal experimentation and testing, is part of what seems to be happening now, which is an increased consciousness of our position in relation to this planet and an increased determination among some of us to hasten this process.

So anyway, back to the chimpanzees. There they are, as I say, ambassadors for the other species and certainly for the other non-human primates. I have had the privilege of working closely with baboons. I spent some time with Rhesus monkeys and I've watched all kinds of different monkeys jumping about in the trees in Africa, and to some extent have learned about their social behaviour. They are very special because they are so like us. How did I leave the forests that I love and start this process towards learning more and more about the use of chimpanzees and other primates in medical research? It actually happened unexpectedly. In 1986 there was a conference in America called "Understanding Chimpanzees" and we brought together all the different people studying chimpanzees in Africa and in some non-invasive captive situations. I went into that conference as a scientist, planning to continue collecting data and analysing it, which I loved. But, we had a session on conditions in captivity for chimpanzees. We had some secretly filmed footage from one of the labs where chimpanzees were imprisoned for medical research. We also had a session on conservation, which showed how the forests with chimpanzees across Africa were vanishing. And I came out of the conference an activist. Since that day in October 1986, I haven't been in any one place for more than three weeks.

I can never forget the shock that I had the first time I went into one of these biomedical labs. It was in Rockville, just outside Washington D.C. I was led first of all into a room where there were these things called isolets – I think they are mostly fazed out now more or less everywhere. These were completely closed with air passing through vents. When the door of one of these isolets was opened, inside was revealed a four-year-old chimpanzee with dull eyes. She was rocking slightly from side to side. She was going to be moved to a larger isolet, because she was outgrowing this one. And so it went on. I went from cage to cage and, by the end of it, I was numb. Even though I'd seen a picture, it was not anything like seeing the reality. Afterwards, I was then taken to a room where there were lot of people from the National Institutes of Health, who were funding this research. We sat around the table and I found that everyone was looking at me. I didn't know what to say, because I was in a state of shock. But then what I did say was: "I am quite sure that everybody in this room, because everyone is a caring, compassionate person, is as shocked as I am." That was kind of clever, because they could not admit that they were not caring or com-

passionate, so they all nodded their heads. And that led to a conference which was funded by HSUS, the Humane Society of the United States of America, where for the first time, we brought together people working on chimpanzees in labs from different labs in the United States and those concerned for their welfare and people from the field. And you know what was strange? Here is another example of how far we've come. And these examples showing how far we have come are very encouraging because, if we have come this far, we can go further. The strange thing was that when I first began sitting down and talking to people who were working with chimps in the labs, there were a lot of animal rights groups who would not speak to me. They said I was talking to the enemy and that we should not do that. I was amazed, because if you don't talk to people, how will you ever move forward? So that conference led to two more.

The first time I saw a fully adult male chimpanzee in a lab I was led up to him by the veterinary. He said, Jojo is very gentle and he won't hurt you, before walking off to do something else. I was wearing a mask, I was made to wear the mask and a little cap and gloves. I knelt down in front of this cage and reached my hand through the bars and Jojo began grooming me through the glove. Suddenly I had this vision of other chimpanzees his age in Gombe, lying on the soft ground, climbing trees, making leafy nests, being groomed by their companions and here was this individual, by himself in a 5-foot by 5-foot steel barred cage where the only sound was the rattling of bars and the screaming of other frightened chimpanzees. I couldn't help it. There were tears running down beneath my mask. Then Jojo reached out one very gentle finger and he lifted my mask and sniffed his finger. It was just the most moving moment, and I think it's moments like that that give you the determination to do whatever you can do about this.

I've continued visiting labs and talking to people, because one of the things I feel very strongly about is that, it's really important that people become truly aware. That they want change not just because it's becoming legislation, not just because people demand different kinds of products and therefore the products have to be made to comply with what people wish, but because people are involved with their hearts. And so my method in dealing with this for me horrific situation has been to try and talk to more and more people.

So just one example of how that can work. I talked about my experience in that first lab I visited. Subsequently, the then very new director of that first lab, John Landon, told me, Jane, at that time I could willingly have murdered you. Even my own daughter heard you speak and was mad at me. But, he said, then I began to think about it and now I really want to thank you, because when I looked at what was going on, I realised it was absolutely inappropriate. And because I realised that, I was able to get money and I'd like to take you back to see the new lab, he said. Well, it's not perfect, far from it, but compared to what I first saw, it must have been like going into some kind of step towards heaven for the chimps.

So getting into peoples' hearts seems to me to be something that is really important. We started a programme called "Roots and Shoots", which is the Jane Goodall Institute's environmental and educational programme for young people. It started in Tanzania 12 years ago and has now spread to 95 countries. It involves projects undertaken by children from pre-school right through university. Roots make a firm foundation. Shoots seem tiny, but many together can break through a brick wall. And if we see the brick wall as all the problems that we have inflicted on this planet, environmental ones, social ones, crime, cruelty, greed and so forth, then the combined hope of hundreds and thousands of young people can break through this brick wall and can make a better world.

I began putting my heart and soul into this program, because I while I was travelling around the world to raise awareness about the plight of chimps, about the plight of forests, about the cruelty, I kept meeting these young people who seemed to have lost hope. When I talked to them, they where either depressed or apathetic or in some cases bitter, angry, even violent, and they all said more or less the same thing. They said, we feel this way because we feel you're generation has compromised our future and there is nothing we can do about it. The world is spinning downward and we feel helpless. So this programme is hands-on action, roll up your sleeves, get out there and do something. And, actually, the very thing that started it was having the chimpanzee as an ambassador, reaching out with one hand to the other animals and with the other hand to us, making the world whole. They work on a project to make things better for their own human community, one for animals, including domestic animals, and one for the environment.

Many of them are deeply concerned about the use of animals in medical experimentation and testing. I have thousands of letters from children. It's desperately important that children learn to understand why things have happened. It's particularly important for them to understand that the people working in these fields are not evil monsters. You'd be amazed how many children automatically assume that anybody who puts on a white coat and does anything with an animal in a lab is some kind of monster. We have to help young people understand what it's all about, where it's come from and where it's going.

Again I would like to come back to what you said, Andrew, because I think we really are at a time now where we can stop thinking about the three Rs and concentrate on one, which is REPLACE. It's not that we can replace everything immediately. You know far more than I do about that. But, work towards it we all can, and I think what should be encouraging here is the number of procedures which can be done without using animals today. Ten years ago people said that would be impossible. So, when I talk to the groups of Roots and Shoots and when I listen to what they've done, when I hear about their peaceful protests outside some of the labs, I say to them, you know, it's just that you need to think differently. Instead of thinking, well, it's really unfortunate that we'll always have to use some animals, but we'll use them as kindly as we can, we'll prevent them from experiencing as much pain as possible, we'll use as few as possible, we'll enrich their lives, we'll try to treat them nicely, but it's sad, we'll always need some. If we turn that around and say, well, think of what we know today about the true nature of animals. Chimpanzees were a bit of a shock to the people back in the early sixties: that they could actually reason, that they could learn sign language and that they could communicate with each

other. This was pretty shocking to a lot of people. But what are we learning today? There are certain things that birds appear to be very intelligent about, but it was always said they couldn't be that intelligent, because their brain did not have the right structure. Thanks to some crows, who are so clever at fashioning tools for specific purposes, people have now re-examined the bird brain and are thinking differently about some of these extraordinary parrots who can learn many, many words. The parrot who knows the most words now has a score of 1350. And no word is counted in his vocabulary unless it's been introduced by him three times in the right context into his sentences, because he speaks in sentences. So, people are now thinking differently about birds. I was just in the most amazing project in Tanzania where giant African rats have been trained to detect landmines. They can smell the explosives. Rats are now also being trained to detect the very early signs of TB from sputum in hospitals. Every rat in that project has a name, every rat's personality is known by the trainer. There is this terrific bond between rat and trainer. We are learning more and more about the tremendously beneficial healing actions of dogs or cats with sick people, we are learning more and more amazing things about rescue dogs, and we have found that autistic and other learning-disabled children have learned to read by talking to a dog or a cat, who is so non-critical, non-judgemental. So, we are learning more and more about animals' brains and abilities and sentience, and we are learning more and more about what they can do for us.

And so it's moving into the time when we need a new mind set. A new mind set that says, given what we know about animals today, it cannot be ethical to put them into tiny cages and do things to them, however nicely we treat them. While we have to, we'll do it and we will treat them as well as possible, but our mindset should be, let's get these extraordinary human brains around the world together and find ways of doing these things without using any animals as quickly as possible.

That's why I was so excited when Andrew said, would I come to speak to all of you, because that's what is happening here. This is the new mindset. That's why you are all here, or most of you. It's very exciting indeed. And with a critical mass of youth growing up to think differently, to understand animals and our role in being good stewards, we'll move further towards that time.

I want to end with two stories. The first is about a chimpanzee, who was captured when he was about two years old. His mother was shot and he was sent to be used for some kind of cancer research. He was named Old Man, because this very small chimpanzee had a wizened face from being sad, as any chimpanzee will be who's had his mother shot and who's been shipped away from Africa. Old Man was lucky, because when he was about 12 years old, they didn't need him anymore for in this work. He and three females were released onto a man-made island at a zoo in Florida and a young man called Mark was employed as a keeper. He was told, don't go anywhere near these chimps, they are much stronger than you are, they hate people, they'll try to kill you. So Mark used to take a little paddle steamer and go over and throw food onto the island. But, he began to watch these chimpanzees and he saw how joyous they were. Every time he appeared with food, before they took a single bite, they would hug each other and kiss each other, make little sounds of happiness, kissing into each others necks. And he thought, how can I care for them if I don't have some kind of relationship with them? A baby was born, Old Man was the father. He loved this baby, he would share food with the child and carry him around and protect him from real and imagined danger. And one day Mark dared to hold out a banana and Old Man took it from his hand and he said to me, Jane, I know just how you felt when David Greybeard took a banana from you! And then the day came when he actually dared to step onto the shore of the island and nothing happened. One day he groomed Old Man and one day they played. They became friends. The females kept away, but they didn't harm him, until one day, when it had been raining, Mark slipped. He fell flat on his face, the baby screamed, the mother acted instantly, as mothers will, rushing in instinctively to protect the child, she bit Mark's neck as he lay on the ground. He felt the blood running down. Then one of the other females rushed up and bit his wrist and one of them bit his thigh. And he thought, how can I ever escape? He looked up and there was Old Man thundering across the island with his hair bristling and his lips bunched in a ferocious scowl, coming to rescue this baby and Mark thought he'd had it. He prepared to die. But what happened? Old Man pulled the three females off Mark and kept them away, screaming and roused as they were, while Mark dragged himself to the boat and to safety. I saw Mark when he came out of hospital, and he said, Jane, there's no question, Old Man saved my life. To me this is so symbolic, because if a chimpanzee, and one who has been abused and harmed by people, can behave thus to reach out across this presumed gap between our species to help a human friend in a time of need, then surely we with our greater capacity for understanding and compassion can do the same for the animals in their time of need?

The second story is about another chimpanzee who was born in Africa. His mother was shot and he was shipped off to a zoo in North America, where he ended up living for about ten years completely on his own. He was named Jojo. He was living in a small old-fashioned cage and then a new zoo director decided to create the biggest exhibit in North America. He built a large enclosure surrounded by a moat filled with water. Of course, chimps don't swim. And he put a number of chimps together into the enclosure including Jojo. Jojo was fine for a while. But then, one of the new young males challenged the senior male. The senior male happened to be Jojo. But Jojo didn't really understand chimp behaviour. He'd been alone and hadn't had the chance to learn. He was terrified by this bristling, swaggering creature displaying at him. And in his fear, he ran into the water. He didn't know about water either. He was so scared that he got over the barrier that was built to prevent the chimps drowning in the deep water beyond. Three times he came up, gasping for air. Then he disappeared under the water. There was just a small group of people on the other side of the moat. There was a keeper, who went off to get a stick to try and scoop Jojo out of the water, but luckily for Jojo, there was a man there who visits the zoo just one day a year with his wife and three little children. He jumped in. He jumped in even though the keeper grabbed him and told him he would be killed. He pulled away. He swam, feeling under the water, until he touched Jojo's body.

He got this 300 pound dead weight over his shoulder, but he felt little movements. Jojo wasn't dead. He got over the barrier and pushed Jojo up onto the bank of the enclosure and then turned to rejoin his rather hysterical family.

There was a woman there with a camera who didn't even remember filming. Most of the video is all over the place, but you can see and hear what happened next. The people on the far side of the moat start screaming at Rick to hurry back and that he's going to be killed. Coming down to see what the commotion is are three big males with bristling hair, and at the same time, Jojo is sliding back into the water, because the bank was too steep. The film amazingly steadies on Rick as he stands there. You see him looking up at his wife and kids, looking up towards where these chimps are coming from, and then you see him looking down at Jojo. For a moment he's completely still. And then he goes back. And again he pushes Jojo up onto the bank. And this time, in spite of everyone yelling at him, in spite of the three chimps approaching, he stays there, he's pushing as hard as he can, and Jojo is making really feeble attempts, trying to grab on to something, and just in time, he grabs hold of a thick tuft of grass and, with Rick pushing, manages to pull himself up to where the ground is level. And, just in time, Rick gets back over that barrier.

That evening that little piece of film was flashed across North America and the then director of the Jane Goodall Institute happened to see it. He called up Rick and said, that was a very brave thing you did. You must have known it was dangerous. Everyone was telling you. What made you do it? Rick said, well, you see, I happened to look into his eyes and it was like looking into the eyes of a man, and the message was, would anybody help me? And, you see, that's the message I've seen in the eyes of little chimps for sale in Africa, from under the frills of the circus, looking out from the 5-foot by 5-foot prisons in the medical research labs, and I've seen it too in the eyes of the monkeys in the labs, and the dogs, and the cats. I've seen it in the eyes of chained elephants, I've seen it in the eyes of children who've seen their parents killed in the ethnic violence in Africa. It's a look that when you see it and feel it in your heart, you have to jump in and try to help.

There's a big challenge ahead of us, for those who want to help, and that is that there is a new calling from the animal protection organisations worldwide to end the use of non-human primates in biomedical research and testing. We are moving in that direction. Some of the big chimp labs have indeed closed. Lemcip closed. Gloria Brazier has 15 of the Lemcip chimps, she knows what they went through. She built a home for them. Some of the others have been housed in different zoos. The big BRPC lab in Holland is closing, they are not working on the chimps at the moment. They have to be housed. The Immunolab that was bought up by Baxter in Austria has closed. Those chimpanzees desperately need a home. The Caustin chimps and the Air Force chimps in New Mexico, fortunately, some white knight sprang to their rescue and with Carolin working to house them, they are moving to Florida to what for them must be paradise. As these labs close, as monkeys get fazed out as well, as they will, the animal welfare community has to be prepared to face a huge challenge. We have to find ways of housing them and caring for them when their days in medical research are ended. Chimp Haven is another sanctuary where chimps from medical research can go, but that's the tip of the iceberg, there are hundreds more chimpanzees needing to be housed and rescued, and thousands of non-human primates around the world.

But if we see that look and feel it in our hearts, we shall rise to that challenge and we shall move towards a time, step by step, sometimes quicker, when we need to use no other-than-human animals in medical research. Because, we do have hearts with compassion, we do have an understanding, we are learning more and more about what it means, the effects of what we are doing to these innocent creatures. So, I have absolutely no doubt at all that, seeing all your faces today, we are moving in that direction much faster than we ever have before. So, let's keep on that tidal wave and inspire more and more people to search for alternatives and not to be afraid of searching for alternatives where it seems that none can ever be. Because sometimes the impossible is possible if enough people get together and use this amazing, extraordinary human brain and decide, we're going to do it, simply because we must.

So, thank you for inviting me here, and for the opportunity, and to all of you for listening and for what all of you are doing, individually or as groups, to help animals. Thank you.

Correspondence to
Jane Goodall, Ph.D.
The Jane Goodall Institute
8700 Georgia Avenue
Silver Spring, MD
USA
www.janegoodall.org
e-mail: cfox@janegoodall.org

Welcoming speech

Reiner Wittkowski
Federal Institute for Risk Assessment, BfR, Berlin, Germany

Minister Künast,
President Tauber,
distinguished guests and colleagues,
and in particular Dr. Jane Goodall

Ladies and Gentlemen,
I am very happy to welcome you here in Berlin to the 5th World Congress on Alternatives and Animal Use in the Life Sciences on behalf of the Federal Institute for Risk Assessment, the BfR.

My colleague President Andreas Hensel has asked me to extend his warmest greetings to you as well.

Some of you may know that being here at this conference in Berlin is like coming back home and being back at the roots.

In November 1986 the Council Directive of the European Union concerning the protection of animals used for experimental and scientific purposes was adopted.

This directive called upon the Commission and the Member States of the European Community to develop and evaluate alternative methods to animal testing as well as to promote research in this field. It was in the context of the adaptation of German legislation on animal protection to this directive – pursuant to which authorisations for animal experiments may no longer be granted if suitable alternative or supplementary methods are available – that the National Centre for Documentation and Evaluation of Alternatives to Animal Experiments, with the abbreviation ZEBET, was founded in 1989 within the former Federal Health Office.

As far as I know, this was the first implementation of a group of researchers following a governmental initiative in the world.

As early as 1991, ZEBET was entrusted with the co-ordination of European validation projects as a result of the competence conceded to it within Europe. Close co-operation exists with the European Centre for the Validation of Alternative Methods (ECVAM), the EU agency which co-ordinates national activities within the European Community and which is meant to strive for recognition of the new methods outside of the European Community.

Absolutely indispensable for the world-wide acceptance of the alternative methods validated in Europe is the co-operation with America's Interagency Co-ordinating Committee for the Validation of Alternative Methods (ICCVAM) which was established almost 10 years ago and represents a total of 14 American federal authorities.

One patent indication of how well this co-operation works is the attendance of 120 US colleagues at today's event.

ZEBET today is integral part of the BfR.

The BfR is the national agency in Germany for regulating the safety of food and of all chemicals except drugs. Research at the BfR is focused on developing test methods for regulatory purposes. At the BfR several national and international reference laboratories are established in the areas of food and chemical safety. It is within the tradition of the BfR to always welcome new and advanced methods to improve consumer safety and also to improve the welfare of animals used in safety tests. Scientists from the BfR are representing Germany in international expert commissions, that are concerned with human safety, e.g. at the EU, OECD and WHO.

The global acceptance of the 3Rs principle "replace, reduce, refine" developed by Bill Russel and Rex Burch almost 50 years ago in the UK proves the humanity criterion, which they proposed in their book "The Principles of Humane Experimental Technique" is finally the benchmark for judging animal experiments from the ethical perspective:

"If we are to use a criterion for choosing experiments to perform, the criterion of humanity is the best we could possibly invent. The greatest scientific experiments have always been the most humane and the most aesthetically attractive, conveying that sense of beauty and elegance which is the essence of science at its most successful."

Ladies and gentlemen, recently four alternative methods had been validated and accepted internationally by the European Union and the OECD.

This is a clear sign that the efforts we all undertake are worth to be undertaken. At present an animal test is under consideration, which I would like to mention exemplarily.

The EU Commission requires for the analysis of marine biotoxins in foods, mussels and shellfish safety testing by the mouse bioassay. The test consists of an injection of homogenates/extracts of mussels into mice to determine the LD_{50}. This test does not only lead to the death of the animals but is in addition extremely painful. Meanwhile, there are alternative methods available based on physico-chemical principles.

I personally welcome that the participants of this congress will discuss this particular problem and help to replace this

regulatory safety test, which is causing pain and suffering and in addition is less sensitive than the new advanced animal free test methods.

Ladies and gentlemen, Horst Spielmann is the head and the spirit of ZEBET from the beginning on and together with his co-workers Manfred Liebsch, Barbara Grune, Richard Vogel and many others they did a tremendous job over years resulting in this wonderful congress. Please allow me to blow my own trumpet now when I explain that I am very proud to have co-workers like them.

But, ladies and gentlemen, scientific expertise and engagement does not suffice to change the world per se. It always needs political backup and support, in particular financial support to attract young scientists to come up with new , methods and strategies. Our Minister Renate Künast is always standing in the first line and fighting when it comes to animal welfare and further improvements to follow the 3Rs principle. Consequently, she has overtaken the patronage of this conference and I cannot imagine that there could be a better forum than this here and today to mention this and say thank you.

Finally, I want to thank all of you for coming to Berlin and participate in a congress that is focusing on an issue that has a high priority at the global level both from the scientific and ethical point of view. I want to thank all of you for devoting your time to the topics of the 5th World Congress and I am looking forward to the results of your discussions, which we will try to implement into our regulatory frame work.

Correspondence to
Prof. Dr. Reiner Wittkowski
Vice President of the
Federal Institute for Risk Assessment BfR
Thielallee 88-92
14195 Berlin
Germany

Special Contributions

ECVAM's Progress in Implementing the 3Rs in Europe

Thomas Hartung
ECVAM – European Centre for the Validation of Alternative Methods, Ispra, Italy

Summary
Starting with the animal welfare directive of 1986 and continuing up to the most recent chemicals and cosmetics legislation, Europe has laid the groundwork for the implementation of alternative methods. In order to meet these political expectations, certain technical and strategic developments became necessary:
- *Analysis of current in vivo test performance to set benchmarks for alternatives*
- *Analysis of the frequency (prevalence) of toxic health effects in different areas of test application*
- *Inventory and database of available alternative methods*
- *Coached development of lacking tests, also making use of novel technologies*
- *Acceleration and international harmonisation of the validation process and regulatory implementation*
- *Development of quality assurance systems for in vitro methods such as Good Laboratory Practice and Good Cell Culture Practice*
- *Transition from single tests as stand-alone replacements to the composition of test strategies and their validation*

The European Centre for the Validation of Alternative Methods (ECVAM) has played a proactive role in all these processes, coordinating many stakeholder activities. A review of the state of these developments shall be given in order to demonstrate how a new type of evidence-based toxicology is emerging, based on validated and quality controlled test strategies.

Keywords: ECVAM, validation, alternatives, political environment

Origin of ECVAM, legal basis, short history

ECVAM is an international reference centre for the development and validation of alternative testing methods aimed at the replacement, reduction or refinement of the use of laboratory animals in the biomedical sciences, with emphasis on toxicological assessments. ECVAM was established by a communication of the European Commission (SEC 91/1794) to the European Parliament and Council referring to Article 7.2 and Article 23 of Directive 86/609/EEC on the protection of animals used for experimental and other scientific purposes. The Directive requires that the Commission and Member States should encourage research into the development and validation of alternative methods that could provide the same level of information as that obtained in experiments using animals, but which involve fewer animals or which entail less painful procedures. ECVAM became operational as a Unit within the EU Joint Research Centre in 1992. ECVAM's work is focused on the development and evaluation of *in vitro* methods (e.g. cell and tissue cultures), of computer modelling based on structure-activity relationships and of physiological and biokinetic modelling. Current political needs for ECVAM's core activities are created by REACH and the 7th amendment to the Cosmetics Directive (Hartung et al., 2003).

ECVAM objectives and strategy

ECVAM pursues its objective to pioneer the process of quality assurance in the life sciences and regulatory testing by:

- communicating regulatory needs to test developers
- coaching the development and optimisation of methods
- tailoring the validation process and good practices
- participating actively in R&D and validation projects
- developing test strategies
- promoting successfully validated tests.

From a strategic point of view ECVAM has worked out a business plan for the next ten years. Overall costs of about € 150 million over these ten years were estimated. These costs will have to be shared with industry, which now urgently needs validated alternatives. Structurally, a dense network of stakeholders is in the process of being established for the development of strategies, the provision of unpublished *in vivo* data and reference chemicals as well as in-house test methods from industry. Candidate tests for validation are selected either in taskforces/ workshops or submitted by test developers, usually via the ECVAM website, and then prioritised by taskforces.

Procedure for validation and regulatory acceptance

ECVAM operates as a coordinator of international validation studies, as a focal point for the exchange of information, as the provider of a central database on alternative methods, as a centre of public dialogue and as a pre-normative research facility of the JRC. Participation in validation studies requires a running infrastructure and active research maintains a practical and realistic view on science and technology. Furthermore, high-quality research ensures credibility in the scientific community.

Due to the political sensitivity of its duties, ECVAM has its own Scientific Advisory Committee (ESAC) composed of members from all 25 European Member States, from relevant industrial associations, from academic toxicology, from the animal welfare movement, as well as from other Commission services with an interest in the area of alternative methods.

ECVAM has established a wide international network with OECD, with its American counterpart ICCVAM and with other Commission services, such as Directorate General (DG) Environment, DG Enterprise, DG Research and Development and DG Health and Consumer Protection. This network is used to reach international expert consent, test implementation and emission of opinions.

A typical validation lasts 3 years (costs per test are about € 300,000) and the subsequent regulatory implementation takes between 2 and 7 years. Main constraints in the process are the availability of reference substances and animal test data, the need for further optimisation of test methods, the duration of financial/administrative procedures and the long-lasting consensus process of regulatory implementation.

Changes in the political environment

The new European legislation on chemicals

The European Commission has proposed to harmonise the testing requirements for existing chemicals, for which there is a lack of safety assessment data (i.e. chemicals marketed before 1981) and new chemicals, by developing a new system for the Registration Evaluation and Authorisation of CHemicals (REACH). The new system will have less stringent testing requirements compared to those imposed by current legislation on new chemicals and will apply to approximately 30,000 substances that are currently marketed in volumes greater than 1 tonne per year. The extent of data requirements will depend on the tonnage of chemical produced in or imported into the EU. Consequently, this will result in a substantial increase of animal use for the safety assessment of chemicals. Several estimates of the number of laboratory animals required for these assessments and the costs for performing the tests have been made. They indicate that several million animals will be required, that testing costs will range in billions of Euros and that the availability of animal testing facilities will be a limiting factor. Beside the ethical aspects and the public concern, economic considerations also call for the timely development and validation of *in vitro* alternatives.

The Seventh Amendment to the Cosmetics Directive

Much of the scientific work on alternatives conducted, coordinated and sponsored in the EU was strongly pushed by the animal protection community and by public opinion, which broadly does not support animal testing for cosmetic products. In the EU, the safety of cosmetic products is regulated by Council Directive 76/768/EEC (EC, 1976). Its 7th amendment was finally approved by the European Parliament and the Council in March 2003. It foresees an immediate ban on animal testing for finished products and a complete ban on animal testing for cosmetic ingredients no later than six years after the implementation of the Directive. Moreover, it requires an immediate marketing ban for new cosmetics (finished products and ingredients) tested on animals where alternative methods validated by ECVAM and accepted by the Community exist. It also foresees a complete marketing ban on cosmetics for some targeted human health effects tested on animals within six years and ten years for repeated-dose toxicities, reproductive toxicity and toxicokinetics. This latter date can be postponed if by that date no validated alternative methods are available.

ECVAM's Strategic Vision

From its inception, ECVAM has been more than the administrator of alternative methods and their formal validation: The field of alternatives requires a proactive contribution (fig. 1), where ECVAM:

a) communicates the regulatory needs to putative developers of new tests
b) surveys opportunities for new technologies
c) steers a strategic debate between stakeholders
d) coaches the development and optimisation of methods
e) develops and tailors the validation process
f) participates in R&D as well as the validation process with its laboratories
g) collects, compiles and provides the information on relevant methods

h) integrates tests into test strategies
i) promotes tests after validation
j) pioneers the process of quality assurance in the life sciences

These roles have to be regularly revisited in light of the political needs. At this moment, two very obvious areas of concern are REACH and the 7th amendment to the Cosmetics Directive. Beside these, even more impressive with regard to animal consumption (fig. 2), are biologicals (16%) when compared to chemicals (1%) and cosmetics (0.025%). Areas like pharmaceutical articles (Hartung, 2002) and basic research (Gruber and Hartung, 2005) also deserve further attention.

Highlights of some recent activities related to the ECVAM strategies include:

a) Communication of needs of the regulatory area to putative developers of new tests
The ECVAM workshop series has just celebrated the 50th workshop report (Gennari et al., 2004) published in the journal Alternatives to Laboratory Animals (*ATLA*). Acknowledging ECVAM's collaboration with FRAME, the publisher of *ATLA*, a joint workshop on "Invalidation of test methods" was carried out in September 2005.

b) Surveying opportunities for new technologies
Following a joint workshop with ICCVAM on validation of toxicogenomics in 2003, a taskforce was established in 2004 and a pilot study was carried out with Affymetrix and Bayer. A workshop on metabolomics in toxicology is in preparation. Very extensive efforts are spent on (Q)SAR (Worth et al., 2004 a & b). This activity – a close collaboration between the European Chemical Bureau (ECB) and ECVAM – was put under the umbrella of OECD, where agreement on the principles of (Q)SAR validation and regulatory use are sought. In parallel, the first formal validation studies were initiated in 2004, addressing skin sensitisation, skin penetration, acute fish toxicity and endocrine disruption.

c) Strategic debate between stakeholders
ECVAM has established a network of about 400 experts regularly working in taskforces and workshops. A broad variety of collaborations with all relevant institutions in the field, aiming to bundle stakeholder activities, also exists. For the first time, a European opinion leader meeting was organised in 2004, convening the organisations acting on the European scale in order to discuss the perception of alternative methods by the (scientific) public and how to improve their image. In close partnership, all ECVAM activities have been opened to the American counterpart ICCVAM. With a view to extend this partnership, the establishment of an International Council of Validation Bodies is under discussion.

d) Coaching of the development and optimisation of methods
The EU has already invested more than € 200 million into the development of alternative methods by funding respective research. Lately, by installing three large Integrated Projects, a new dimension of tailored development of alternatives was reached (fig. 3). These projects, which involve more than 90 institutions and a funding of about € 40 million aim to make available batteries of tests plus the respective test strategies within five years each.

ReProTect (Hareng et al., 2005), started in July 2004, and deals with the field of reproductive toxicology including endocrine disruption. Noteworthy, the different tests included in the project for optimisation and integration into a test strategy originate not only from the field of alternatives, but also from areas such as mechanistic biomedicine, especially reproductive medicine, pharmaceutical agent discovery, clinical diagnostics, breeding of farm animals, etc. These models were never suggested as alternative methods, since they have different purposes and reflect only partial aspects of the reproductive cycle, but, put together in a conceptual framework, they might allow building a predictive test strategy.

A-Cute-Tox was based on an ECVAM workshop in 2003 (Gennari et al., 2004). It started in January 2005 and aims to establish an animal-free classification of acute toxicity, substituting for the classical LD_{50} test. Several studies have shown a good correlation of *in vitro* cytoxicity studies with the animal experiment. The project aims to improve this correlation to an acceptable level by outlier reduction.

Sens-it-iv was started in November 2005 with a view to complete the development of animal-free test strategies for skin and

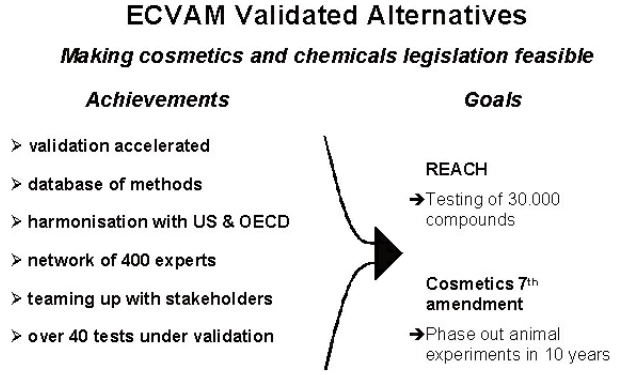


Fig. 1: ECVAM validated alternatives

Purposes of animal experiments in Europe in 2002

Total number	**10,700,000**	**100%**
Safety evaluations	**1,060,000**	**10 %**
Agricultural chemicals	**123, 000**	**1 %**
Industrial chemicals	**136,000**	**1 %**
Cosmetics	**2,700**	**0.025%**

Fig. 2: Purposes of animal experiments in Europe in 2002

lung sensitisation. It was based on an ECVAM workshop (Casati et al., 2005).

e) Development and tailoring of the validation process
On the one hand just recently for the first time international consensus on the role and procedure of the validation process was reached by the creation of OECD guidance document 34 on the validation and international acceptance of new or updated test methods for hazard assessment. The guidance document mainly reflects the ECVAM principles. On the other hand, various challenges to the validation process as such exist. For example, the enormous investments involved for the validation of a single test with regard to laboratory work and costs have often been questioned. In light of the need to validate a very high number of tests for the purposes of the chemicals and cosmetics legislations this aspect had to be reviewed.

So far, validation studies did not make use of existing data. However, in some instances, such retrospective validation might be an appropriate shortcut to the assessment of validity. Further challenges to the current validation scheme originate from new technologies (pattern-based or "-omics" approaches, transgenic animals, *in silico* methods such as (Q)SAR or computer modelling). In response, ECVAM has proposed a modular approach (Hartung et al., 2004), which opens up ways to accommodate these needs. The discussion as to the optimisation of the validation procedure, however, continues.

f) Participation of ECVAM in R&D as well as in the validation process with its laboratories
As part of the EU Joint Research Centre, ECVAM is also a place of research and education. Repeatedly, ECVAM has made contributions to the development and optimisation of alternative methods. ECVAM's laboratories also allow active participation in validation studies. Details are given in a chapter below.

g) Collection, compilation and provision of information on relevant methods
ECVAM hosts a database on alternative methods (dbAlm), which provides high-quality information related to alternative methods, including protocols for relevant *in vitro* tests. The major part of this scientific information service will be available online by the end of the year. A valuable resource of documents, such as the ECVAM workshop reports, is provided free of charge via the ECVAM website (www.ecvam.jrc.it).

Another major contribution to the field has been the compilation of inventories of alternative methods that are currently available. An inventory of 280 pages compiled with 75 experts in the context of the cosmetics legislation was published as an *ATLA supplement*.

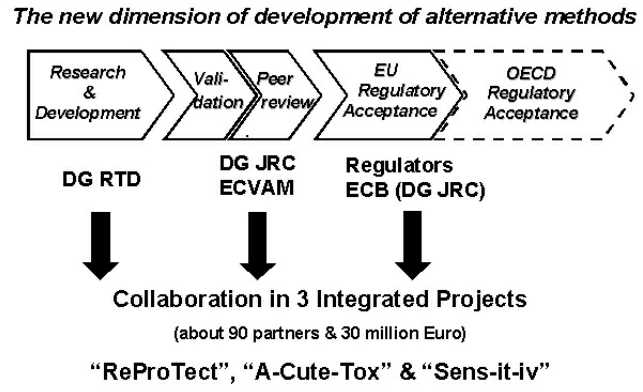


Fig. 3: The new dimension of development of alternative methods

h) Integration of tests into test strategies
Many of the more complex toxicological endpoints will not be replaced by single alternative methods. Instead, it will be necessary to develop testing strategies based on batteries of tests and their intelligent combination in test strategies. An important element of this is the concept of prevalence of health effects of chemicals (Hoffmann and Hartung, 2005), i.e. the actual proportion of chemicals showing a certain toxic property. Furthermore, it is necessary in order to develop such strategies, to analyse the performance of the animal experiment, an effort which has just started for example in the field of skin irritation (Hoffmann et al., 2005). Methods adapted from decision theory and evidence-based medicine will be employed in order to compose and validate final testing strategies.

i) Promotion of tests after validation
Today the regulatory implementation of validated tests often lasts longer than the validation process itself. This obvious bottleneck can only be overcome by collaboration with regulators, such as the National Coordinators of the Chemical Test Guideline Program in Europe, or on the level of the OECD. However, for this purpose, validation has to take into account the needs of regulators, i.e. a validity statement is less a scientific judgement but more a proof that the method is fit for its (regulatory) purpose.

j) Pioneering the process of quality assurance in the life sciences
Further to an ECVAM workshop in 1999, an OECD guidance document on Good Laboratory Practice (GLP) for *in vitro* toxicological studies was accepted in 2004 (fig. 4). In parallel, a Good Cell Culture Practice (GCCP) guidance document (Hartung et al., 2003) was completed recently (Coecke et al. 2005). It sets the minimal standards for quality control of *in vitro* work and will enable an international discussion over the next year. In the context of quality control, ECVAM is aiming for further adoption of principles on evidence-based medicine to the field of toxicology. The establishment of a taskforce on evidence-based toxicology was agreed in 2004.

Validated and accepted alternative methods

Eight alternative methods for chemicals/cosmetics have been endorsed by ECVAM including tests for skin corrosivity, skin sensitisation, phototoxicity as well as embryotoxicity. Seven alternative methods reached scientific acceptance for e.g. the safety evaluation of biologicals such as vaccines. Eleven alternative methods for myelotoxicity, pyrogen testing, acute fish

toxicity, mutagenicity and chronic toxicity in dogs are currently under peer-review by ESAC.

Relevant methods were accepted by both the European Commission (Annex V to Council Directive 67/548/EEC on the classification, labelling and packaging of dangerous substances) as well as by the Test Guideline Programme of the OECD. Three potency tests of biologicals have been accepted by the European Directorate for the Quality of Medicines (European Pharmacopoeia).

Estimated impact of current activities

The impact of alternative methods can be very substantial: The most successful replacement method, the Limulus test for pyrogens, reaches an annual turnover of € 200 million and saves more than one million rabbits per year. The remaining 200,000 rabbits used per year in Europe can most probably be fully replaced by the five methods currently being peer-reviewed (Hoffmann et al., 2005). For the time being, a reduction method saving one third of animals was proposed (Hoffmann et al., 2005).

For chemicals and cosmetic ingredients the following animal experiments (% animal use in Europe in 1999) have to be considered:

- acute toxicity (35%): OECD-accepted alternative methods in 2000 reduced the number of animals used from 45 to 8 per chemical; alternative methods which have been currently validated (to be completed 2005) reduce this number again to an estimated 3-6 animals per chemical. ECVAM set up the Integrated Project A-Cute-Tox (2005-2009, 9 million € funding, 37 partners) which aims for full replacement of the animal tests based on an ECVAM workshop (Gennari et al., 2004).
- Skin sensitisation (5%): The OECD-accepted refinement method has reduced the number of animals used per chemical as well as their suffering. ECVAM set up an application for an Integrated Project Sens-it-iv (2006-2010, € 11 million funding, 31 partners) following an ECVAM workshop (Casati et al., 2005).
- Chronic toxicity (27%): a strategy for chronic toxicity was developed at an ECVAM workshop held in 2004; currently a consortium and a work programme are being set up.
- Toxicokinetics (2%): a strategy for toxicokinetics was developed at an ECVAM workshop in 2004. The OECD accepted a test for skin penetration in 2004. First validations on barrier models (blood brain barrier and gut absorption) and work on physiology-based pharmacokinetic modelling as well as metabolism start in 2005.
- Mutagenicity and carcinogenicity (8%, costs of an animal cancer study are as high as 800,000 Euro per substance): two validation studies started in 2005 in which variants of the cell transformation assay and the *in vitro* micronucleus test are being evaluated.
- Reproductive toxicity (13% of animal use in toxicology in 1999; up to 55% of the testing costs of REACH): Three embryotoxicity tests were validated in 2002; ECVAM set up the Integrated Project ReProTect (2004-2008, 9 million €, 27 partners) to develop an alternative test strategy.
- Endocrine disrupters (2%): Validation studies with the US were agreed on for 2006 and will be part of ReProTect.
- Skin-eye corrosion (3%): OECD accepted tests in 2004.
- Phototoxicity (3%): OECD accepted test in 2004.

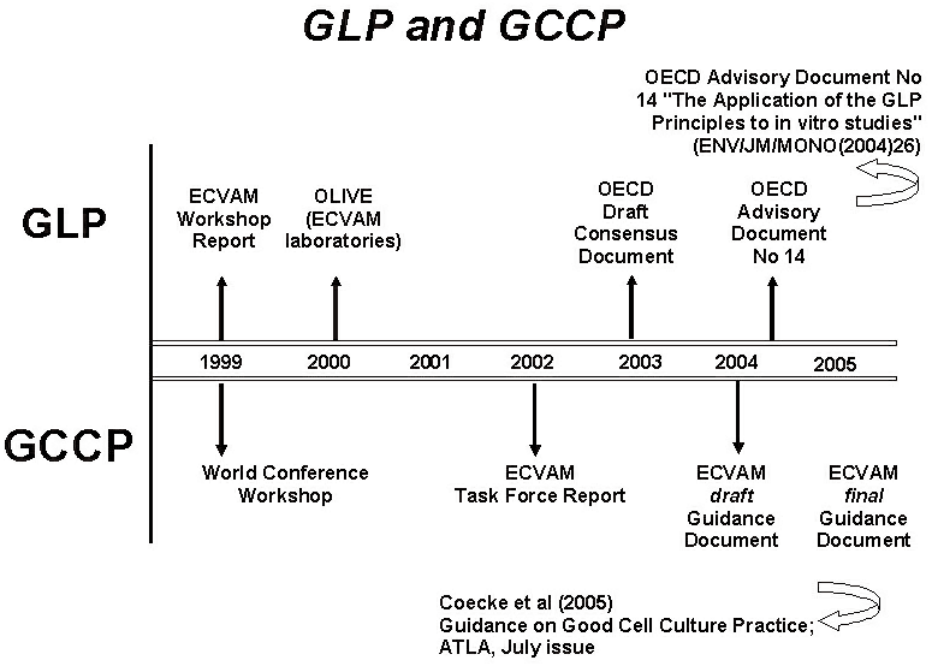


Fig. 4: GLP and GCCP

- Skin irritation (1%): a validation study is currently underway (2004-2005).
- Eye irritation (1%): Joint review of existing data with US ICCVAM from 2004-2005. Depending on the outcome, joint validation studies will be organised in 2006-2007.
- Acute fish toxicity (up to 40% of animals used for REACH): The reduction alternative developed by ECVAM will reduce the number of fish used by 60%. The method will be peer-reviewed in 2005. The fish egg test, which is a full replacement of the current animal test, will be validated in 2006-2007.

The impact of QSAR (*in silico* methods) cannot yet be anticipated, since no experience with validation exists. Regulatory use so far was mainly in priority setting and not in substitution of the current animal tests. Furthermore, it is not clear which percentage of chemicals can be subjected to QSAR (problems with lack of purity, mixtures, salts, metal compounds, etc.). Calculations by the ECB suggest an up to 60% reduction. Similarly, possible use of toxicogenomics is encouraged.

Beside the aspect of animal reduction, cost savings which, in some cases, can reach more than 90%, and a higher throughput, which will be extremely relevant for the testing of 30,000 chemicals under REACH, will have to be considered.

ECVAM's own research activities

Participation in validation studies requires a running infrastructure. The active participation of ECVAM in validation studies, which ranges from checking SOPs and transferability to full participation in blinded ring trials under GLP, helps identify problems of methods and allows flexible filling of gaps in studies. It also ensures neutrality of the study. Since GLP is increasingly becoming standard for validation studies, ECVAM is currently establishing this quality assurance regimen.

Active research maintains a practical and realistic view on science and technology. ECVAM must not become a purely administrative body, because the validation process is always interlinked with developmental aspects. The pipeline of methods to be validated has to be filled actively in interaction with the basic and applied research community. ECVAM's own research safeguards that this dialogue remains realistic and effective.

High-quality research also ensures credibility in the scientific community. This requires visibility and, most importantly, publications, favourably in higher impact factor journals. Currently, ECVAM researchers publish about 40 scientific papers per year, also in prestigious journals such as Nature and the Proceedings of the National Academy of Science.

ECVAM's laboratories have high standards with regard to infrastructure, space and resources. A very unusual combination of technologies and expertise (e.g. various *in vitro* technologies, metal toxicology, stem cells) is further amplified by the links to neighbouring units and the enormous network of external collaborators. This provides an excellent environment for about 20 Ph.D. students, post-docs and visiting scientists. Altogether, a unique integrated approach spanning from basic to applied research and the regulatory view is possible. ECVAM's research is uniquely positioned between basic research and validation.

Given the difficult access to animal primary cells at ECVAM and the limitations of cell lines, clear priority is given to human (primary) cells. The emerging stem cell technologies (embryonic and adult) (Bremer and Hartung, 2004; Pellizzer et al., 2005) offer new opportunities as do the classic accessible human cell sources blood and bone marrow. Cryopreserved human blood (Schindler et al., 2004) represents a very promising source of standardised cell material without the problem of blood donations. Among the animal cell lines, Balb 3T3 deserve special attention, since they are used both for the cell transformation assay and for the currently validated basic cytotoxicity test.

Several projects have led to the identification of prototypic toxins, e.g. metal compounds or test substances from validation studies. In some instances, ECVAM possesses high-quality *in vivo* and human data. Many of these substances have already been characterised in several standardised *in vitro* systems. This allows synergy linking between projects by testing the same set of reference substances. This will be further increased by the planned establishment of repositories of substances. Currently, a high-throughput testing facility is established jointly as a JRC exploratory research project, which offers opportunities for several projects, such as those on acute toxicity (A-Cute-Tox), acute fish toxicity, neurotoxicity and immunotoxicity.

Work on human (stem) cells will always be limited by the number of cells available. Furthermore, organotypic (co)cultures require analysing individual cells in mixtures, which is also the case for most stem cell derived differentiated cells. This calls for the establishment of technologies allowing single cell analysis, such as confocal microscopy, FACS, cell sorting (all existing at ECVAM), laser scanning microscopy, *in situ* PCR, laser microdissection, cell chips, etc. Setting up an array of cutting-edge methods should keep ECVAM attractive for visiting scientists.

As another aspect, the fate of test substances *in vitro*, e.g. their solubility, binding to plastic or serum albumin, etc., is routinely not considered. Addressing this might improve the predictive capacity by reducing an uncertainty factor. Similarly, the effect of exposure patterns to test substances *in vitro* has hardly been studied. Several current projects aim to understand the effective exposure of cells *in vitro* and the consequences for toxic responses.

Signature-based ("-omics") and computational models promise new approaches in several fields. The full integration in all key areas will be important to leverage these technologies. Current projects make use of toxicogenomics and metabolomics (Nuclear Magnetic Resonance, NMR, and Mass Spectroscopy, MS).

Co-operation between ECVAM and USA

Interactions between ECVAM and governmental bodies in the USA started as early as 1993, shortly after the creation of ECVAM. Since 1995, ECVAM has had a bilateral co-operation with the US Interagency Co-ordinating Committee on the Validation of Alternative Methods (ICCVAM). Its aim is to

ensure an early exchange of information on the validation of test methods to facilitate mutual recognition, acceptance and implementation of scientifically validated testing methods. Secondly, this co-operation serves to facilitate the OECD process in providing harmonised protocols to the scientific community and promoting international adoption of validated alternative methods.

The existing collaboration between ECVAM and ICCVAM in the field of alternative testing methods has been strengthened during the last three years and comprises the following activities: ICCVAM has an observer status on the ECVAM Scientific Advisory Committee. The Head of ECVAM became member of SACATM, the US Scientific Advisory Committee for Alternative Toxicological Methods. Both ESAC and ICCVAM have agreed on parallel peer-review and arbitration of results for the upcoming peer-reviews (pyrogen tests, haematotoxicity, chronic toxicity in dogs, micronucleus test). ICCVAM and ECVAM have agreed on creating an International Council of Validation Bodies to coordinate validation studies at the level of OECD. Discussion about formal collaboration with OECD has been initiated. About 20 visits of ICCVAM members or ICCVAM-nominated experts to ECVAM taskforces, workshops and validation management groups take place per year. The FDA has allotted a specific budget for parts of these travel costs. A sabbatical programme to exchange ECVAM and ICCVAM personnel was agreed upon. In 2003, the Head of ECVAM also became member of the Scientific Advisory Committee of CAAT, the Center for Alternatives to Animal Testing at Johns Hopkins University, Baltimore, which has pioneered the field of alternative methods in the US for about 25 years. At the same time, he became member of the Scientific Advisory Committee of the Institute for In-Vitro Sciences (IIVS), Gaithersburg. Furthermore, a senior American manager from The Procter and Gamble Company is on secondment for two years at ECVAM.

Good Laboratory Practice – Good Cell Culture Practice

The requirement for carrying out validation studies under standardised conditions, i.e. GLP and GCCP rules, has been recognised by national and international validation bodies. ECVAM plays a leading role in this process and actively contributes to the drafting of advisory and guidance documents. ECVAM, DG Enterprise and ICCVAM were part of a GLP working group which drafted an OECD Guidance document on GLP and *in vitro* toxicology that was finalised in May 2004.

ECVAM is also playing a leading role in drafting a new Guidance Document on Good Cell Culture Practice (GCCP). The aim of this GCCP document is to reduce uncertainty in the development and application of animal and human cell and tissue culture procedures and products by encouraging greater international harmonisation, rationalisation and standardisation of laboratory practices, quality control systems, safety procedures, recording, reporting and compliance with regulations and ethical principles. In order to give this document an international dimension, ECVAM invited ICCVAM to be part of the steering group that drafted this document (Coecke et al., 2005).

Databases

As an outcome of a project of the ECVAM Task Force on Alternatives Databases in collaboration with the Head of the thesaurus section of the US National Library of Medicine (NLM), a first version of a thesaurus on animal alternatives has been developed using the novel "bottom-up" approach. The thesaurus was generated in a semi-automatic manner, by selecting actual phrases that occurred in 2000 documents, and should therefore reflect the preferred terminology used by the authors of the articles. This first version focuses on toxicity testing and the first 11 main sectors identified. Following two consultation rounds, the thesaurus will be made available throughout the new Internet version of the ECVAM Database for Alternative Methods (formerly ECVAM Scientific Information Service) with practical application for end-users expected in fall 2005.

ECVAM has been invited to participate in a newly formed ad hoc group of the NLM to address the following main subjects:

- Journals on Alternatives and MEDLINE
- Keywords on alternatives in MEDLINE
- Adding Search Filters to PubMed

The following significant outcomes have been reached: Based on the advice given by the participants during the expert meeting, the Directorate of the NLM decided to add 9 journals on alternatives to MEDLINE, change the index terms for MeSH to better identify papers related to the alternatives concept and to sponsor a study on the feasibility of adding query filters for animal alternatives searches.

References

Bremer, S. and Hartung ,T. (2004). The use of embryonic stem cells for regulatory developmental toxicity testing in vitro – the current status of test development. *Curr. Pharm. Des. 10*, 2733-2747.

Casati, S., Aeby, P., Basketter, D. A. et al (2005). Dendritic cells as a tool for the predictive identification of skin sensitisation hazard. *ATLA – Altern. Lab. Anim. 33*, 47-62.

Coecke, S., Eskes, C., Gartlon, J. et al. (2005). The value of alternative testing for neurotoxicity in the context of regulatory needs. *Environmental Toxicology and Pharmacology 33*, 261-287.

Coecke, S., Balls, M., Bowe, G. et al. (2005). Guidance on Good Cell Culture Practice. *ATLA – Altern. Lab. Anim. 33*, 261-287.

Gennari, A., Berghe, C., Casati, S. et al. (2004). Strategies to Replace In Vivo Acute Systemic Toxicity Testing. *ATLA – Altern. Lab. Anim. 32*, 437-459.

Gruber, F. P. and Hartung, T. (2004). Alternatives to animal experimentation in basic research. *ALTEX 21 Suppl. 1*, 3-31.

Hareng, L., Pellizzer, C., Bremer, S. et al. (2005). The Integrated Project ReProTect: a novel approach in reproductive toxicity hazard assessment. *Reproductive Toxicology 20*, 441-452.

Hartung, T., Bremer, S., Casati, S. et al. (2004). A Modular Approach to the ECVAM Principles on Test Validity. *ATLA – Altern. Lab. Anim. 32*, 467-472.

Hartung, T. and Gribaldo, L. (2004). New hepatocytes for toxi-

cology? *Trends Biotechnol. 22*, 613-615.

Hartung, T., Bremer, S., Casati, S. et al. (2003). ECVAM's response to the changing political environment for alternatives: Consequences of the European Union chemicals and cosmetics policy. *ATLA – Altern. Lab. Anim. 31*, 473-481.

Hartung, T. (2002). Three Rs Potential in the Development and Quality Control of Pharmaceuticals, *ALTEX 18 (Suppl. 1)*, 3-11.

Hartung, T., Balls, M., Bardouille, C.et al. (2002). Report of ECVAM task force on good cell culture practice (GCCP). *ATLA 30*, 407-414.

Hoffmann, S., Cole, T. and Hartung, T. (2005). Skin irritation: prevalence, variability, and regulatory classification of existing in vivo data from industrial chemicals. *Regul. Toxicol. Pharmacol. 41*, 159-166.

Hoffmann, S. and Hartung, T. (2005). Diagnosis: Toxic! – Trying to apply approaches of clinical diagnostics and prevalence in toxicology considerations. *Tox. Sci. 85*, 422-428.

Hoffmann, S., Luderitz-Puche, l U., Montag-Lessing, T. and Hartung, T. (2005). Optimisation of pyrogen testing in parenterals according to different pharmacopoeias by probabilistic modelling, *J. Endotoxin Res. 11*, 25-31.

Hoffmann, S., Peterbauer, A., Schindler, S. et al. (2005). International validation of novel pyrogen tests based on the human fever reaction. *J. Immunol. Meth. 298*, 161-173.

Pellizzer, C., Bremer, S. and Hartung, T. (2005). Developmental toxicity testing from animal towards embryonic stem cells. *ALTEX 22*, 47-57.

Schindler, S., Asmus, S., von Aulock, S. et al. (2004). Cryopreservation of human whole blood for pyrogenicity testing. *J. Immunol. Meth. 294, 89-100.*

Worth, A. P., Hartung, T. and Van Leeuwen, C. J. (2004a). The role of the European centre for the validation of alternative methods (ECVAM) in the validation of (Q)SARs. SAR QSAR *Environ. Res. 15,* 345-358.

Worth, A. P., Van Leeuwen, C. J. and Hartung, T. (2004b). The prospects for using (Q)SARs in a changing political environment--high expectations and a key role for the European Commission's joint research centre. *SAR QSAR Environ. Res. 15*, 331-343.

Acknowledgement

The contributions of ECVAM senior staff to parts of the text are gratefully appreciated.

Correspondence to

Prof. Thomas Hartung, MD, PhD
European Commission, JRC – Joint Research Centre
Institute for Health and Consumer Protection
ECVAM – European Centre for the Validation of Alternative Methods
Via E. Fermi 1, TP 580
21020 Ispra (VA)
Italy
e-mail: Thomas.HARTUNG@cec.eu.int

The Three Rs: Looking Back ... and Forward

Michael Balls
FRAME, Nottingham, UK

Summary
Personal experiences at FRAME and ECVAM are recalled, alongside the evolutionary progress of the Three Rs (reduction, refinement, replacement) approach to animal experimentation, described in great decades analogous to the Great Ages of Western Civilisation.

Keywords: animal experimentation, ECVAM, FRAME, reduction, refinement, replacement

Introduction

When I was asked to prepare this talk for a session entitled *Looking Back – The Voices of Experience* at the Berlin Congress, I decided to revisit the great ages of the Three Rs, originally presented in my plenary lecture at the 1996 Utrecht Congress as my interpretation of the development of alternatives over decades in analogy to the Great Ages in the evolution of Western civilisation over centuries (Balls, 1997), but this time placing greater emphasis on my own experiences during the last 50 years.

The 1950s: The age of renaissance

The 1950s saw the birth of the Three Rs (*reduction, refinement, replacement*) concept, as a result of the vision of Charles Hume of UFAW and the penetrative thinking and skilful analysis of William Russell and Rex Burch, culminating in the publication of *The Principles of Humane Experimental Technique* (Russell and Burch, 1959).

I was at high school at the time, majoring in biology, where I was better at botany than zoology, but could never remember the names of the many different plants. As a result, apparently being sufficiently familiar with the comparative anatomy of the rat, frog and dogfish, as well as knowing enough Latin, I found myself reading zoology at Oxford.

The 1960s: The age of darkness

Despite the warm welcome given by the Scientific Establishment to the Three Rs concept, little seemed to happen during the 1960s, and Russell and Burch had gone off to further their careers in other ways. However, at the end of the decade, Dorothy Hegarty and Charles Foister, impatient with the squabbling between the pro- and anti-vivisectionists, founded an organisation with a more positive goal, the Fund for the Replacement of Animals in Medical Experiments – FRAME (Annett, 1995). I have searched through all the FRAME archives, and I can find no evidence that the two founders of FRAME had ever heard of the Three Rs, so I truly think that this was a case of what we zoologists would call convergent evolution, i.e. ending up at the same place via different routes.

When preparing for my finals in 1960, my ambition was to be a zoologist with the British whaling fleet in the Antarctic, or, if I had to do compulsory military service, an officer on the bridge of a destroyer carving its way through the Atlantic. Instead, Michael Fischberg asked me to join his experimental embryology group at Oxford, where my project was to transplant cancer cell nuclei into *Xenopus laevis* eggs, to see whether the egg cytoplasm could reprogramme the errant genome to produce a normal embryo.

To do this exciting work (which turned out to be technically impossible, by the way), I had to work on amphibian tumours, and this meant having a vivisection licence. These were the days when it was politically acceptable to use words like "vivisection" – now we have to talk about "animal procedures". Getting a licence under the *Cruelty to Animals Act 1876* required the signatures of the head of a Royal college and of a medical professor. I therefore went to see Professor Sir Lyndor Brown, Professor of Physiology at Oxford, with a carefully worded draft of what I proposed to put on the licence application form. "No", he said, taking a small piece of paper from a drawer in his desk, "that won't do at all. Use these words and you'll be able to do whatever you like." I did, and got my licence.

Shortly afterwards, Fischberg was offered chairs at Edmonton, Geneva and Yale, and asked me to go with him, whichever choice he made. As a result, I had three very happy years in Geneva, before going as a post-doc to the University of California at Berkeley, where I learned cell culture with Harry Rubin, in the research group where Howard Temin discovered reverse transcriptase, later being awarded a Nobel Prize. I then went on to Reed College in Portland, Oregon, to join Larry Ruben, with whom I have continued to study various aspects of the biology of amphibians, ever since we first met in Geneva in 1962.

By 1966, I was back in the UK as a lecturer in developmental biology at the University of East Anglia. In 1967, a young man came into my office and said that, since he had just become a biology teacher at a local high school, he wondered if he could

do some part-time research. This was Richard Clothier, and our friendship and partnership continue to this day.

The 1970s: The age of reason

Interest in the issues raised by animal experimentation increased during the 1970s, especially in Britain, where it focused on the centenary of the *Cruelty to Animals Act 1876* and the need to reform it. Meanwhile, Andrew Rowan was Scientific Director at FRAME, and in 1978, he organised a meeting at the Royal Society on *The Use of Alternatives in Drug Research* (Rowan and Stratman, 1980), one of the first meetings on the application of replacement, at which I was one of the speakers.

Richard Clothier and I had been trying to work with our students on the cell cycle and the control of cell division in amphibian tissues *in vitro*. Along the way, we discovered that tissues such as liver, kidney and pancreas from *Amphiuma means*, the Congo eel (not from the Congo, but from the Southern USA, and not an eel, but a newt-like amphibian), could survive as organotypic cultures for several weeks, whereas their mammalian equivalents could only last for a few hours. Unfortunately, there was little sign of any cell division, so, desperate to help our students to get the results they needed for their Ph.D. theses, we started working on tissue functions instead, along the way becoming comparative endocrinologists, pathologists, pharmacologists, physiologists and toxicologists, as well as comparative anatomists.

To get closer to medical matters, we moved to the University of Nottingham Medical School in 1975, shortly after which we were visited by David Smyth, who was conducting a survey for the Research Defence Society, which was to lead to the publication of *Alternatives to Animal Experiments* (Smyth, 1978). He asked us if we knew that we were working on alternatives, to which we replied in the negative, and he suggested that we should try to get research support from animal welfare organisations. We took up this suggestion and were delighted to get some significant help from the Humane Research Trust.

Meanwhile, Andrew Rowan persuaded Dorothy Hegarty that I might make a good FRAME Trustee, so, in 1979, I was invited to join FRAME. He had also suggested that FRAME should set up an independent committee to look at the application of the Three Rs in toxicology and toxicity testing. The Trustees put me on the committee to watch over FRAME's interests, whereupon I was made Chairman, since I was the only member who had no right to call himself a toxicologist.

The 1980s: The age of reformation

At the end of the 1970s, Merlyn Rees MP, Home Secretary in the then Labour Government, who was responsible for the regulation of animal experimentation and who had come under great pressure to reform the 1876 Act, said that he would not meet lots of different groups of campaigners separately, but would be prepared to consider an agreed joint submission. This led to the formation of an alliance between the British Veterinary Association (BVA), the Committee for the Reform of Animal Experimentation (CRAE) and FRAME, which produced its proposals in 1983 (Anon, 1983). By this time, there had been a change to a Conservative Government, and David Mellor MP was the Home Office minister given the task of preparing and introducing the new legislation.

The BVA/CRAE/FRAME proposals were used as the basis for what was to become the *Animals (Scientific Procedures)* Act 1986, and representatives of the Triple Alliance advised the Government at every stage, from the drafting of the Bill to the Royal Assent. We succeeded in securing a number of major inclusions, such as the establishment of an independent Animal Procedures Committee (APC), and the requirement that the Home Secretary must weigh the balance between likely (animal) suffering and likely (human) benefit, before granting a project licence.

It was while serving on the APC that I first came across Russell and Burch's book. There was no copy at FRAME, so Clive Hollands lent me his, which I photocopied. Nearly 15 years later, my friend, Rodger Curren, presented me with a first edition of *The Principles* during the Bologna Congress, which Bill Russell signed there and then.

Changes were also taking pace in Europe, which saw the introduction of Three Rs legislation in the form of *Directive 86/609/EEC* and Council of Europe *Convention ETS123*, and in the USA, with the passing of the *Animal Welfare Act*.

These were also exciting times at FRAME. I had become Chairman of the Trustees on the retirement of Mrs Hegarty, and moved the FRAME headquarters to Nottingham. *ATLA Abstracts* was re-launched as a typeset journal, now called *ATLA (Alternatives to Laboratory Animals)* and with an international editorial board, and the report of the Report Toxicity Committee was published and discussed at a conference held at the Royal Society. Thanks to David Mellor, long a friend of FRAME and now a patron, we received the first grant ever given by British Government specifically to support alternatives. We used it to establish the INVITTOX database (which is still in use, but is now owned and run by ECVAM), and the FRAME International Alternatives Validation Scheme (which established principles which were later to be the basis of the ECVAM/ICCVAM/OECD principles that are in force today). We also began collaborative research with the University of Nottingham and others, with the support of various industrial companies, which led to the establishment of a FRAME Alternatives Laboratory (FAL) in the Medical School, as well as to the development of the kenacid blue cytotoxicity test.

The 1990s: The age of revolution?

My plenary lecture was given in 1996, and was aimed at challenging the participants of the Utrecht Congress to join me as revolutionaries – hence the question mark. I will return to that in due course.

There had already been a very successful congress at Baltimore in 1993, and the CAAT/ERGATT and ERGATT/EC workshops on the validation and regulatory acceptance of alter-

native toxicity test methods had established sound foundations on which a peaceful revolution could be built in that field.

My own life had changed considerably. First, I had become Professor of Medical Cell Biology at the University of Nottingham, and the Vice-Chancellor had given me his explicit support in my role as *de facto* honorary Director of FRAME, which meant that I spent almost all my time at the FRAME Office, while Richard Clothier ran the FAL.

I had expected to spend the rest of my professional and personal life in Nottingham, but I was persuaded to accept an invitation to apply to become the first Head of the European Centre for the Validation of Alternative Methods (ECVAM), which was being set up as part of the Environment Institute at the European Commission's Joint Research Centre, at Ispra, near Lago Maggiore, in north-western Italy. This came about because, having (I thought anonymously) advised a London agency during the preparation of a report for the Commission's Environment Directorate General (then DGXI), I had been rewarded by being asked to do various other small jobs for DGXI, including advising on the possibility of establishing a validation centre (I advised against it!).

Thus, offered the position on 15 March 1993, by 2 April I was already at ECVAM. Remarkably, I somehow managed to retain my roles as Chairman of the FRAME Trustees and Editor of *ATLA* throughout my stay in Italy, and, in the days before e-mail, I was able to keep contact with Nottingham via the fax machine in our villa, overlooking Lago di Monate. I take great pleasure from the fact that FRAME was going from strength to strength following the appointment of a real Director, Robert Combes, while the FAL continued to flourish under Richard Clothier.

Reminiscences of ECVAM (1993 – 2002)

Becoming a civil servant in a rigid hierarchical system was a great shock to me, as I had never before had to receive or obey orders from my superiors. There was betting in Brussels that I wouldn't survive for more than six months, and I was told, on more than one occasion, that I was to be sacked and that my successor had already been appointed. However, my immediate boss, Fritz Geiss, Director of the Environment Institute, advised me to ignore all that, decide what needed to be done, and get on and do it well. I took this advice, and stayed in place while three Commissioners, four Director-Generals and five Directors came and went. I have many amazing stories to tell, but they will have to wait until I publish my memoirs, if I ever summon up the courage to do so.

I now look back on this period with much pleasure and with great gratitude – living in Italy (the scenery, the culture, the food and wine, the people) was an absolute joy, leading an international team of gifted young people was a great privilege, and developing ECVAM's vast international network of collaborators was vitally important. Chairing the ECVAM Scientific Advisory Committee (ESAC) was both stressful and enjoyable, and I can trace most of what ECVAM did under my leadership to what was discussed with the members of the ESAC.

When I went to ECVAM, I said that there would be only one true measure of our success (or failure) – the number of alternative methods validated as reliable and relevant for their particular purposes. We had agreed with DG Environment that our target would be 15 methods by the end of the 5th Framework Programme in 2003, and I am proud of the fact that 16 methods had been endorsed by the ESAC when I retired (at a time of my own choosing, by the way) in 2002 (Balls, 2002). In addition, ECVAM assisted in securing the regulatory acceptance of *in vitro* methods for percutaneous absorption, and also of three refinement acute *in vivo* toxicity tests, which permitted the deletion of the OECD LD_{50} test guideline.

I feel that I must list just a few of many other ECVAM highlights, such as:

1. The ECVAM workshop series, based on the successful CAAT/ERGATT/EC model – a series set up soon after I arrived at Ispra, because ECVAM had to spend a lot of money quickly, or lose it (Combes, 2002). Many of these workshops were held at the Hotel Lido, Angera, where much of the controversy and bitterness of the day dissipated as we ate and drank excellent Italian food and wine as the sun set on the other side of Lago Maggiore. In September 2005, an ECVAM/FRAME workshop was held, partly to celebrate the success of the series of more than 50 workshops held so far.

2. *The Three Rs: The Way Forward* workshop, held in Norfolk about 1 km from where we now live, and organised by Bill Annett of FRAME, and jointly chaired by Alan Goldberg for CAAT, and me for ECVAM (Balls et al., 1995). This was the first meeting that Russell and Burch had attended together since 1959, and sadly, it was to be the last, as Rex Burch, already very ill, died a few months later. The meeting was held in Norfolk because of his illness, beginning with an opening ceremony in Sheringham Town Hall, where he had his laboratories. A frequent visitor to the town, I had passed within a few metres of him on many occasions over about 30 years, but I hadn't noticed the name plaque on the wall!

3. In addition to the workshop series, ECVAM also had a number of task forces, i.e. small groups of individuals asked to do specific tasks. One of the most important of these was the prevalidation task force, which in a few days produced an historic document by putting the meat on the bones of a prevalidation scheme which Rodger Curren and I had roughed out while having a coffee overlooking Baltimore harbour (Curren et al., 1995).

4. Having me as Editor of *ATLA* also had advantages for ECVAM, since I was able to get things published within weeks. One significant occasion was when this ensured that the ECVAM validation principles (Balls and Karcher, 1995) could be published just before the ICCVAM and OECD held meetings on validation, which eventually led to what are now known as the ECVAM/ICCVAM/OECD principles.

5. One of the most memorable events of the 1990s was the *3rd World Congress on Alternatives and Animal Use in the Life*

Sciences, held in Bologna in 1999. The Congress was organised by ECVAM, with Marlies Halder playing the anchor role, and with invaluable support from the Public Relations Unit of the Joint Research Centre and FRAME.

The 2000s: The age of achievement ...

The current decade is one of great challenge and great opportunity for alternatives, not least in Europe, because of the EU REACH system for new and existing chemicals and the 7th Amendment to the EU Cosmetics Directive, and the resultant unparalleled support for ECVAM from the European Commission and the European Parliament.

I retired from ECVAM in June 2002, but I am continuing to support ECVAM and my friend and successor, Thomas Hartung, mainly through editing and publishing ECVAM reports of various kinds, as well as participating in workshops. Meanwhile, I am privileged to be able to continue to support FRAME, especially through working with Bob Combes, Gerard Duvé and our gifted and energetic young colleagues.

That the 5th Congress in Berlin has more than 850 participants is an indication that much is being achieved, but the question is, "Is it enough, and will the revolution that didn't occur in the 1990s, now take place in the 2000s?" Sadly, the question mark I used in 1996 was entirely justified.

or of disappointment?

I see a number of disturbing trends, which cause me great concern.

The progressive *reduction* in the numbers of animal experiments, which had been foreseen when the new legislation was passed in the 1980s, seems to have come to an end, especially as more and more mice are sacrificed on the altar of genetic exploitation. Also, far from working together toward the zero option of the use of non-human primates, there is pressure to build more and more primate research centres. *Refinement*, claimed by many to be the poor relation among the Three Rs, seems to be becoming increasingly fashionable, especially as it can be linked with the indefinite continuation of the reliance of research and testing on animal models. However, for me at least, refinement cannot answer many of the fundamental questions. Giving a monkey a tennis ball to play with, and hiding its food so it must search for it, may be better than long confinement in a barren cage, but that's not enough for me. This is in line with the thinking of Russell and Burch, who said: *Refinement is never enough, and we should always seek further reduction and if possible replacement. … Replacement is always a satisfactory answer*. (Russell and Burch, 1959, p. 66).

Therefore, our goal must be *replacement*, as Jane Goodall and Andrew Rowan emphasised at the opening lecture in this Congress. Achieving it will require skill, dedication and commitment – and the energy of young people trained in modern approaches to medical problems and toxicity testing. We have begun the journey, but there is still a very long way to go.

So, my message as I look to the future is the same as that I gave at the end of the Bologna Congress, when I wore a tie given to me by my friend, Klaus Cussler, of the Paul Ehrlich Institute, Langen, Germany (Balls, 2000). It has about 100 tortoises on it, all moving slowly in the same direction. But one of them is saying to the others, "GET A MOVE ON!"

Please, dear readers, don't disappoint me.

References

Annett, B. (1995). The Fund for the Replacement of Animals in Medical Experiments (FRAME): The first 25 years. *ATLA 23*, 19-32.

Anon. (1983). *Animal Experimentation in the United Kingdom: Proposals Submitted to the Home Secretary Jointly by the BVA, CRAE and FRAME*. Edinburgh: CRAE.

Balls, M. (1997). The Three Rs concept of alternatives to animal experimentation. In L.F.M. van Zutphen and M. Balls (eds.), *Animal Alternatives, Welfare and Ethics* (27-41). Amsterdam: Elsevier.

Balls, M. (2000). The reduction, refinement and replacement of laboratory animal procedures – a look to the future. In M. Balls, A-M. van Zeller and M.E. Halder (eds.), *Progress in the Reduction, Refinement and Replacement of Animal Experimentation* (1763-1765). Amsterdam: Elsevier.

Balls, M. (2002). The establishment of ECVAM and its progress since 1993. *ATLA 30, Suppl. 2*, 5-11.

Balls, M., Goldberg, A. M., Fentem, J. H. et al. (1995). The Three Rs: The way forward. *ATLA 23*, 838-866.

Balls, M. and Karcher, W. (1995). The validation of alternative test methods. *ATLA 23*, 884-886.

Combes, R. D. (2002). The ECVAM workshops: a critical assessment of their impact on the development, validation and acceptance of alternative methods. *ATLA 30, Suppl. 2*, 151-165.

Curren, R.D., Southee, J.A., Spielmann, H. et al. (1995). The role pf prevalidation in the development, validation and acceptance of alternative methods. *ATLA 23*, 211-217.

Rowan, A.N. and Stratman, C. J., eds. (1980). *The Use of Alternatives in Drug Research*. London: Macmillan.

Russell, W.M.S. and Burch, R. L. (1959). *The Principles of Humane Experimental Technique*. London: Methuen.

Smyth, D. (1978). *Alternatives to Animal Experiments*. London: Scolar Press.

Correspondence to

Prof. Michael Balls
FRAME
Russell & Burch House
96-98 North Sherwood Street
Nottingham NG1 4EE
UK
e-mail: Michael.Balls@btopenworld.com

Theme 1
Education

Session 1.1
Refinement and reduction alternatives in education: Teaching humane science

Online Learning to Teach Humane Science

Nicole E. Duffee
American Association for Laboratory Animal Science, Memphis, USA

Summary
The AALAS Learning Library (ALL) is a learning management platform that offers laboratory animal science courseware with documentation of training. Online learning in the laboratory animal field supports the 2Rs of reduction and refinement in animal use, and in addition, the ALL specifically teaches the 3Rs in courses where ethics is addressed. With now 103 courses, curricula are in development for users with technical, research, IACUC (ethical committee), management, training, and veterinary functions. Users are mainly in the USA, and international users from many countries worldwide are a growing faction.

Keywords: online learning, courses, courseware, training documentation, learning management

Introduction

Education and training programs for animal researchers focus on reducing the use of animals in research and on refining animal handling and treatment when animal use is necessary. As an adjunct to face-to-face training of research personnel, online learning provides depth in knowledge of concepts, prepares a learner for personal training, and reinforces lessons learned.

Options are increasing for the online education and training of scientists on the ethics of animal research. An online learning management system (or platform) has been developed by the American Association for Laboratory Animal Science (AALAS Learning Library at www.aalaslearninglibrary.org) and the U.S. Department of Veterans Affairs (www.researchtraining.org). Since the release of this platform in 2001, over 50,000 US researchers have completed ethics courses, and access continues to grow monthly at the rate of about 2,000 individuals. Currently, 457 institutions use this platform to train all research staff and 52 institutions use it to train a core group of animal facility staff.

This paper focuses on the progress made in the development of the AALAS Learning Library (ALL) in which AALAS has undertaken an initiative to expand the course curriculum of the AALAS Learning Library (ALL) to fully support the 3R's, particularly the R's of reduction and refinement, via promoting the competence of personnel in animal research. AALAS is developing curricula on the ALL for five functions for research personnel: research, technical, management and training, veterinary, and institutional animal care and use committee. Building competence with an emphasis on bioethics in each area of staff function will enhance animal welfare in research.

Platform

The ALL platform (fig. 1) is a learning management database in SQL Server that provides interactive courseware, transcript data for training activities, and automatic systems for enrolment, account purchase and renewal, user management, and author and administrative interfaces.

Course materials include text in html, interactivity via javascript, images, streaming media, practice questions, and exams. Multiple formats of questions are available for practice quizzes and exams. Feedback text may be incorporated in questions answered either correctly or incorrectly. Training activity is tracked by access of each page in a course, completion of a course, and passing of exams. Continuing education units (CEUs) are awarded on the basis of the completion of a course and the passing of the corresponding exam. Users may print a certificate on demand for CEUs earned.

Selected courses are free (*Working with the IACUC, IACUC Essentials, and Working with the Laboratory Mouse*, as well as some certification courses), but most are accessible only with an active account. An account provides access for one year to all courses including new ones launched during the term of the account. This approach of opening wide the door to every course through low fees and unrestricted access on the ALL lowers the barriers to receiving education and training.

To accommodate different sizes/types of institutions with varying needs, the ALL architecture provides access for both individuals and groups. A group may be an institution or an AALAS Branch organisation. Prices are reduced in individual accounts for AALAS members ($100 USD) and in group accounts for AALAS institutional/commercial members (lowest cost is $10 USD per person in a group of 150 or more persons).

In addition to the individual and group accounts, the ALL is structured for creating a custom library, in which specific courses may be selected for a custom curriculum and their materials may be modified to more closely support the needs of an institution's animal welfare program. The custom library allows an institution to educate and train large numbers of research staff combining both AALAS and institution-specific content. AALAS course offerings can be tailored to fit an institution's needs by adding or deleting material or images. (Only the certification courses are exempted from this capability since their content is tied to a suite of certification exams.) Moreover, through a simple authoring interface, institutional staff can both modify an existing course or create a new one. With a custom library set up for an institution, staff log in and enter their own institution's course library. The custom library model is intended help an institution address either short-term critical and long-term global training requirements. In 2004, for example, a university requested a custom library to meet their short-term goal in rapidly training 4,500 researchers on animal regulatory topics. Authorised individuals at the university customised the *Working with the IACUC* course, making it specific to their institution's policies and procedures, and this course was placed in a customised library that only university personnel could access. With the placement of that one course in their custom library and the distribution of login instructions to all personnel, the university staff navigated easily to the correct course and completed it without confusion.

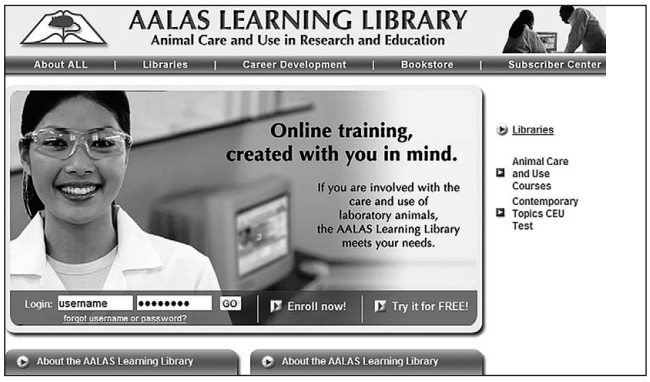


Fig. 1: AALAS Learning Library home page
(http://www.aalaslearninglibrary.org)

The ALL also incorporates automatic features to support administrative functions for users. A group has a Co-ordinator with authorities to manage the group's members, i.e., enrolling group members, assigning accounts to members, detaching members no longer associated with the group, viewing member transcripts, and downloading reports on member transcripts as Excel files or csv, txt, or html files. Users who purchase an account (Group Co-ordinator or individual users) receive automatic renewal notices and may view order history.

Any user may subscribe to the ALL listserv to receive announcements about the site, new courses, and information on how best to use the ALL platform.

Outcomes

Users

The ALL has 52 groups with 2,536 accounts and 80 individual users with accounts. Additional users who access the free courses without enrolling are not tallied. Although most users are based in the USA, 126 users are in countries in Europe, Latin America, the Middle East, Asia, and Oceania. For example, 6 New Zealanders recently obtained accounts through their membership of the AALAS Branch "Palms to Pines", which is located in California. Two institutions in Singapore enrolled their staff (a total of 21 persons) in group accounts to help them prepare for AALAS technician certification. The availability of the courseware 24 hours a day, 7 days a week, makes the ALL convenient to use in all world time zones. For persons outside the USA who are interested in attaining AALAS certification, the ALL offers the most economical way to study for all three examination levels. And, although no international group has a custom library at this time, a custom library would allow the production of courses on regulatory affairs and issues which are relevant to a country or region.

Courses

The AALAS Online Learning Committee oversees the curricula for the AALAS Learning Library (ALL) and has defined curricular goals for five functions ("training tracks") of animal facility/laboratory personnel: Research, Technical, Veterinary, Management and Training, and Institutional Animal Care and Use Committee (IACUC). These tracks serve to organise the course topics by staffing function. Because a subscriber may access any training track to take any course desired, the track feature accommodates persons with multiple duties (e.g., veterinary medical care, management, and technical procedures). In each track, the curriculum has objectives of imparting technical competence and ethical guidance on the use of animals in research. Research and IACUC courses feature US regulations, ethical decision-making, minimising pain and distress, and concepts of analgesia and anaesthesia. Additional research courses have topics of mouse bioengineering, mouse breeding, animal

use methodologies, and facility operations. Certification-related courses prepare technicians to ascend the AALAS certification ladder, which provides qualifications to work with research animals and which is recognised by USA regulatory authorities. In the Management and Training Track, courses for trainers aim to impart skills for improving how to teach staff and test for knowledge retention. Management courses address facility operations plus human resource/supervisory skills. Veterinary courses relate to the field of laboratory animal medicine. Table 1 arranges series of courses by training track, and table 2 lists all individual courses available. Depending on the nature of the content, some courses overlap categories of staff function and so appear in multiple training tracks (tab. 1), whereas other courses are unique to a staff category and so are listed in only one corresponding training track. The Online Learning Committee members and AALAS staff will continue to produce courses in all categories.

The approach toward course development has been a broad-based effort to include experts in the laboratory animal field to co-author and/or review course materials. Contributors have joined these efforts via responses to listserv invitations for subject matter experts, by referrals following a presentation or paper of topical interest to a course in planning, and by general networking. As of August 2005, 132 individuals participated in the development of the ALL courses through contributions of content, questions, images, and media.

Ethics, and specifically the 3Rs, are repeatedly cited in courses that address animal use so that responsibilities toward animals are couched in these terms. An example of these is *Ethical Decision-Making in Animal Research* (Technical, Research, IACUC, and Management & Training training tracks). This course, developed with support by the Charles River Laboratories Foundation, teaches how to consciously think through a process for reaching an ethical decision in challenging situations in a research setting, such as whether to euthanise an animal prior to the end of a study. Distinctions are made among a range of outcomes that are ethically required, permitted, encouraged, and prohibited. Six steps are described in this approach to:

- Identify the ethical question raised in the case.
- Gather and assess all relevant facts.
- Identify the stakeholders.
- Identify the values that play a role in the decision.
- Identify possible solutions.
- Test out the decision by thinking what would follow if this decision were publicly known to be allowed in these kinds of cases.

Two related courses in the "Case Studies – a Trainer's Resource" series offer case studies on ethical situations so one may practice working through the ethical issues of real-life situations in animal research. The case studies provide a worksheet for entering notes about each step of the ethical analysis. Case study printouts may be used for leading group discussions on the ethical dilemmas in research. Available within each case study course is a trainer's resource of handouts and a guide which aid a group discussion on the case study.

The course and case studies were developed by Deni Elliott, PhD (the Poynter-Jamison Chair of Media Ethics and Press Policy at the University of South Florida.); Judy Murray, BA (Technical Training Coordinator, Charles River Laboratories); Stacy Marco, MBA, MS, LATG (Operations Manager, MD Anderson Cancer Center); Sally Walshaw, VMD (Director of Animal Resources, Atlantic Veterinary College); and Nicole Duffee, DVM, PhD.

User surveys

In December 2003 (5 months after launching the AALAS Learning Library) and again in September 2004, the AALAS Online Learning Committee invited all enrollees to take part in online questionnaires for user information and satisfaction.

In the 2003 questionnaire (78 respondents), most users (82%) said that they enrolled on the ALL to study for certification exams (36%), obtain continuing education (28%), or do their job better (18%). Users (82%) thought that it was easy to enroll in the ALL. Most users rated the courses as good (58%) or excellent (38%). When asked about what they liked best about the ALL, 28% liked the year-long access to every course, 31% liked the continuing education certificates, and 26% liked the transcript documentation.

In the 2004 survey (147 respondents), users again indicated that they were highly satisfied with the AALAS Learning Library: 77% reported that the ALL platform was easy to use, and 20% reported that access or navigation problems were rare. Most users (99%) rated the courses as good (40%) or excellent (59%). This survey provided a glimpse into the respondents' educational backgrounds and job responsibilities showing the diversity among the individuals who use the ALL. Most users had either a bachelor's degree (37%) or a high school diploma or GED (23%), and there were users as well with associate's, master's, and doctoral degrees. Most users worked as animal care technicians (26%), and many were also research technicians (17%), veterinary technicians (11%), facility managers (15%), scientists (6%), and facility directors/administrators (9%). Users expressed their preferences for courses mainly in regulatory and IACUC related topics (40%), as well as certification (28%), pain management (15%), and occupational health and safety (15%).

In each survey, open-ended questions allowed respondents to provide feedback and recommendations on course topics and platform issues. In response to specific comments, improvements were implemented wherever possible.

Conclusion

AALAS is committed to providing education on animal welfare and the 3Rs through all its educational resources and publications. The AALAS Learning Library extends that goal into online training in the laboratory animal field and supports the Rs of reduction and refinements in animal use. Knowledge in ethical and technical concepts in animal research is essential for both developing an animal protocol with the appropriate number

of animals and conducting procedures which minimise the induction of animal pain and distress. Through the dissemination of online education and training, the AALAS Learning Library promotes the responsible care and use of animals in research to benefit people and animals.

Correspondence to
Nicole E. Duffee, DVM, PhD
American Association for Laboratory Animal Science
9190 Crestwyn Hills Dr.
TN 38125 Memphis, USA
e-mail: nicole.duffee@aalas.org

Tab. 1: Course and exam titles by training track on the AALAS Learning Library.
(The number of courses and exams are shown when there are multiples per title.)

Course and exam titles (Number of courses, exams if multiple per series)	Training tracks				
	Technical	Research	IACUC	Management & Training	Veterinary
Assistant Laboratory Animal Technician (29 courses, 33 exams)	÷				
Laboratory Animal Technician (14 courses, 18 exams)	÷				
Laboratory Animal Technologist (18 courses, 22exams)	÷				
Taking an AALAS Technician Certification Examination	÷				
Animal Welfare Act Regulations	÷	÷	÷	÷	
Public Health Service Policy on Humane Care and Use of Laboratory Animals	÷	÷	÷	÷	
Guide to the Care and Use of Laboratory Animals	÷	÷	÷	÷	
Good Laboratory Practice Standards	÷	÷	÷	÷	
Ethical Decision-Making in Animal Research (3 courses/exams)	÷	÷	÷	÷	
Pain Recognition and Alleviation in Laboratory Animals	÷	÷	÷		
Post-Procedure Care of Mice and Rats in Research: Minimising Pain and Distress	÷	÷	÷		
Working with the IACUC		÷	÷		
Writing a Protocol for Research in Animals (9 courses/exams)		÷	÷		
Essentials for IACUC Members			÷		
Introduction to Ergonomics in the Laboratory Animal Facility	÷			÷	
Ergonomics for Animal Technicians: Working Smart	÷			÷	
Biosafety in Microbiologic and Biomedical Laboratories	÷	÷	÷	÷	
Health Risks and Safety Procedures for Working with Nonhuman Primates	÷	÷	÷	÷	
Video: Working Safely with Nonhuman Primates*	÷	÷			
Working with the Laboratory Mouse	÷	÷			
Aseptic Rodent Surgery	÷	÷			÷
Mouse Breeding Colony Management	÷	÷		÷	
Genetically Engineered Mice: Historical Perspectives	÷	÷			
Genetically Engineered Mice: Approaches for Evaluating the Phenotype	÷	÷			
Transgenesis and Conditional Control Systems	÷	÷			
Biosecurity Issues Related to Genetically Engineered Mice	÷	÷			
Basic Metrics for the Laboratory Animal Facility	÷				
Selection of Cage Cleaner Products - Chemistry Driven				÷	
Workplace Training				÷	
Writing Multiple Choice Questions				÷	
Active Listening				÷	
Team Building				÷	
Time Management and Goal Setting				÷	
Introduction to Laboratory Animal Medicine					÷
Animal Observations and Clinical Signs of Disease	÷				÷
TOTAL Courses and Exams (103/109):	85/91	28/28	22/22	18/18	3/3

Tab. 2: Courses available on the AALAS Learning Library

Assistant Laboratory Animal Technician Level Courses
- ALAT 1: History and Purpose of Laboratory Animal Science and Animal Care Programs*
- ALAT 2: The Research Facility Environment*
- ALAT 3: An Introduction to Science
- ALAT 4: Cell and Tissue Structure
- ALAT 5: Organs and Organ Systems
- ALAT 6: Feed and Nutrition
- ALAT Review Exam (1-6)‡
- ALAT 7: Heredity and Breeding
- ALAT 8: Laboratory Animal Environment
- ALAT 9: Facility Equipment
- ALAT 10: Hygiene in the Laboratory Animal Facility
- ALAT 11: Animal Procurement
- ALAT 12: Health and Disease
- ALAT 13: Drug Therapy and Common Diseases of Laboratory Animals
- ALAT 14: Euthanasia
- ALAT 15: Experimental Design and Methodology
- ALAT Review Exam (7-15) ‡
- ALAT 16: Mice
- ALAT 17: Rats
- ALAT 18: Hamsters
- ALAT 19: Guinea Pigs
- ALAT 20: Gerbils
- ALAT 21: Rabbits
- ALAT 22: Cats
- ALAT 23: Dogs
- ALAT 24: Nonhuman Primates
- ALAT 25: Swine
- ALAT 26: Sheep/Goats
- ALAT 27: Amphibians
- ALAT 28: Birds
- ALAT 29: Miscellaneous Laboratory Animals
- ALAT Review Exam (16-29) ‡
- ALAT Final Review Exam (1-29) ‡

Laboratory Animal Technician Level Courses
- LAT 1: Public and Private Interests in Animal Research*
- LAT 2: Administrative Responsibilities*
- LAT 3: Laboratory Techniques
- LAT 4: Genetics and Breeding
- LAT 5: Anatomy and Physiology
- LAT Review Exam (1-5) ‡
- LAT 6: Laboratory Animal Facility Equipment
- LAT 7: The Laboratory Environment
- LAT 8: Animal Health Maintenance
- LAT 9: Health and Disease
- LAT 10: Diagnostic Techniques
- LAT Review Exam (6-10) ‡
- LAT 11: Aseptic Technique, Surgical Support and Anesthesia
- LAT 12: Emergency Veterinary Care
- LAT 13: Research Methodology
- LAT 14: Calculations and Conversions
- LAT Review Exam (11-14) ‡
- LAT Final Review Exam (1-14) ‡

Laboratory Animal Technologist Level Courses
- LATG 1: Functions of Management*
- LATG 2: Identifying and Controlling Costs*
- LATG 3: Regulations and Security
- LATG 4: Quality Assurance
- LATG 5: Occupational Health & Safety
- LATG 6: Structure and Function of Cells and Tissues
- LATG 7: Organic Chemistry and Biochemistry
- LATG 8: Molecular Biology
- LATG 9: Genetic Engineering
- LATG Review Exam (1-9) ‡
- LATG 10: Infectious Diseases
- LATG 11: Immunology
- LATG 12: Common Diseases of Laboratory Animals
- LATG 13: Diagnostic Techniques
- LATG 14: Pharmacology
- LATG 15: Anesthesia and Analgesia
- LATG 16: Unique Anatomical Features of Some Laboratory Species
- LATG 17: Gnotobiology
- LATG 18: Statistics
- LATG Review Exam (10-18) ‡
- LATG Review Exam (1-18) ‡
- Taking an AALAS Technician Certification Examination‡

US Federal Regulations and Guidelines
- Animal Welfare Act Regulations
- Public Health Service Policy on Humane Care and Use of Laboratory Animals
- Guide to the Care and Use of Laboratory Animals
- Good Laboratory Practice Standards

Bioethics
- Ethical Decision-Making in Animal Research
- Ethical Case 1: Mouse in a Parasitology Experiment
- Ethical Case 2: Rat with Partial Paralysis

Pain and Distress Management
- Pain Recognition and Alleviation in Laboratory Animals
- Post-Procedure Care of Mice and Rats in Research: Minimizing Pain and Distress*

IACUC Courses
- Working with the IACUC: VA version*
- Working with the IACUC: non-VA version*
- Essentials for IACUC Members*

Writing Animal Protocols
- Writing an Animal Protocol for Research on Mice*
- Writing an Animal Protocol for Research on Rats*
- Writing an Animal Protocol for Research on Rabbits*
- Writing an Animal Protocol for Research on Guinea Pigs*
- Writing an Animal Protocol for Research on Hamsters*
- Writing an Animal Protocol for Research on Gerbils*
- Writing an Animal Protocol for Research on Dogs*
- Writing an Animal Protocol for Research on Cats*

Occupational Health & Safety
- Introduction to Ergonomics in the Laboratory Animal Facility
- Ergonomics for Animal Technicians: Working Smart
- Biosafety in Microbiologic and Biomedical Laboratories
- Health Risks and Safety Procedures for Working with Nonhuman Primates
- Video: Working Safely with Nonhuman Primates*

Mouse Biomethodology
- Working with the Laboratory Mouse*
- Mouse Breeding Colony Management
- Genetically Engineered Mice: Historical Perspectives
- Genetically Engineered Mice: Evaluating the Phenotype
- Transgenesis and Conditional Control Systems
- Aseptic Technique for Rodent Survival Surgery

Math
- Basic Metrics for the Laboratory Animal Facility

Cage Washing
- Selection of Cage Cleaner Products - Chemistry Driven

Management & Training
- Writing Multiple Choice Questions
- Workplace Training
- Active Listening
- Team Building
- Time Management and Goal Setting

* Course can be taken without having an account, but there is no transcript documentation.

‡ These courses are designed to help individuals prepare for the certification exams, they do not provide CEUs.

Three Barriers Obstructing Mainstreaming Alternatives in K-12 Education

Lynette A. Hart, Mary W. Wood and Hsin-Yi Weng
UC Center for Animal Alternatives, School of Veterinary Medicine, University of California, Davis, USA

Summary
The limited use of alternatives in secondary education contrasts with the concerted adoption of alternatives in veterinary curricula. Taking a teacher's perspective, three barriers obstruct mainstreaming of alternatives in high school biology courses. First, dissection is not addressed in course outlines, curricular standards, and frameworks. Second, financial and technical support for resources in science teaching is lacking. Third, teachers need ways to motivate their students to learn biology and offer them stimulating, informative materials. Preparation of appropriate materials for five to ten biology laboratories could address these three barriers at modest cost and effectively deliver biology to secondary students.

Keywords: biology, secondary education, dissection, teaching, alternatives, intermediate education, physiology, systems

Introduction

Using dissection or experimentation with animals as a method of teaching students has a colourful and contentious history (Tansey, 1998; Klestinec, 2004). Ideas in science education have evolved through the nineteenth and twentieth centuries (DeBoer, 1991), yet the use of animals in laboratories of secondary schools in the United States has continued with little change or educational scrutiny.

Considering the uses of animals in research, teaching, and testing, the uses in education seem most amenable to replacement, the most sought after of the 3Rs. Indeed, veterinary schools increasingly have mainstreamed alternatives in their curricula and a large number of teaching resources are available (Hart et al., 2005), but widespread adoption of alternatives has not yet occurred for teaching high school biology in the United States. Monitoring the use of animals in education, while not a comprehensive effort, indicates there has been a sharp reduction of animal use in higher, medical, and veterinary medical education, but perhaps less reduction at the secondary level. Having a substantial replacement of animal use in the advanced education and training of veterinarians, but still using many animal specimens in high school teaching, looms as a growing paradox that continues as a subject of criticism and controversy.

Consistent with the adoption of alternatives in higher education, an ever-growing supply of resources exist, almost 4,000, that are cataloged in the NORINA database (Smith, 2005). Some of these resources are categorised and described in the InterNICHE book and website (Jukes and Chiuia, 2003; InterNICHE, 2005), as well as the AVAR website (AVAR, 2005). Both traditional and alternative resources, including dissection materials and various models, are advertised widely by distributors (Carolina Biological Supply Company, 2005; NASCO Online Catalogs, 2005). Despite this great number of resources, secondary school teachers still are not offered and provided a well-integrated package of resources that interface with and complement the curricular lessons for courses in high school biology and physiology.

Over the years, papers and books concerning opposition to use of animals in education have presented relevant analyses, including reviews of ethical considerations for alternatives (Langley, 1991), and the patterns of use in the United States (Orlans, 1991), in other countries (Balcombe 2000a), and in higher education (Balcombe 2000b). Reflecting the controversial nature of the topic, in some papers, the posture has been frankly political or philosophical, arguing that we should or should not allow dissection (Sapontzis, 1995; Kline, 1995). Much of the controversy concerning animal use in secondary education has focused directly on communication with students, providing legislated protection to those who prefer to use alternatives and coaching them in strategies to avoid dissection (Balcombe, 1997a). Another topic, again focusing on students, has been considering the adverse ethical consequences of instructing students to be involved in harming or killing animals (Orlans, 2000).

A study from England reported on a survey of 468 students regarding their experiences with and attitudes toward animals in education, including dissection (Lock and Millett, 1992). A subsequent survey assessed the use of animals from the teacher's perspective (Adkins and Lock, 1994). About a third of the teachers held opinions that discouraged them from using animals in their teaching. In an Australian study, all 34 surveyed schools reported doing dissection, limited primarily by cost, and almost all schools also included activities with living animals (Smith, 1994). A retrospective study in Canada sought to document students' experiences and attitudes, both positive and negative, concerning dissections they performed in secondary school (Bowd, 1993). Lock (1994) replied, agreeing with many aspects of Bowd's paper, but differed in having the view that no alternatives were superior to dissection. An ethnographic study was

conducted later to learn more about the reactions of students to their experiences in dissection (Barr and Herzog, 2000). Like Bowd, they found that a substantial minority viewed dissection primarily in negative terms.

Comparing the performance and achievement of high school biology students who use simulated dissection versus actual dissection, the simulation was equally effective for learning (Kinzie et al., 1993). A review of various studies using simulated alternatives for teaching anatomy, at various academic levels, also found that simulations yielded similar achievement outcomes as live dissection, whether using low-tech or high-tech simulations (Zirkel and Zirkel, 1997). Yet another approach has been to use the simulation as a preparation for the dissection, resulting in the students learning more anatomy following the dissection (Akpan and Andre, 2000). In general, the issue of dissection appears to have been most visible during the late 1980s and early 1990s, then giving way in the educational community to emphases on standards, curricula assessment, and diversity. Dissection is strikingly absent from published materials on standards and frameworks.

With these laboratories, there may be a gap between the objectives set for the laboratory and the accomplished outcomes associated with the expected learning (Ralph, 1996). A recent assessment of laboratories in U.S. high school science curricula by the National Research Council has concluded that, in general, the quality of current laboratory experiences is poor for most students, and that improving high school science teachers' capacity to lead laboratory experiences effectively is critical (Singer et al., 2005). Additional criticism was levelled at the organisation and structure of most high schools, the state science standards, and the current large-scale assessments. Similar criticisms were reported from an earlier study of the laboratory work in British Columbia High Schools, a report that called for substantial research and reform (Gardiner and Farragher, 1999).

These varied perspectives have not considered the constraints teachers face, but rather have criticised teachers. Balcombe (1997b) evaluated some of the barriers against acceptance of alternatives in teaching, including that some teachers are resistant to change; it requires investing time and money; information on alternatives is not widely disseminated; and the quality of material available varies. In more recent writings Balcombe (2001) has directly made a case for adoption of alternatives rather than using dissection. In this presentation, we build on the paper by Balcombe (1997b) concerning the barriers against acceptance of alternatives. Taking the teachers' perspective, we propose three barriers mitigating against rapid adoption of alternatives in classrooms, recommending production of web-based teaching resources to address these barriers and improve instruction in biology laboratories, especially within the United States.

Methods

Two groups of pre-college teachers participated in discussions that contributed to this paper. A group of 23 teachers worked with us during the academic year, 1993-1994, using instructional software, "The Virtual Heart", in their classrooms (Zasloff and Hart, 1997). During 2003-2004, 5 teachers participated in a focus group and subsequently continued as consultants in further discussions. The teachers all were teaching in public junior or senior high schools in the Sacramento Valley during the period of their participation.

Barriers against adoption of alternatives

Teachers generally are highly motivated to employ the best teaching materials and resources they can feasibly acquire, often even purchasing materials with their own funds. Although the topics of biology that lend themselves to dissection are interesting to students, the curriculum is very full with information required to be taught, leaving little laboratory time for most teachers when teaching mammalian biology. Commonly, about five laboratory sessions are scheduled, sometimes as double periods. Occasionally, a semester-long physiology course is offered, permitting more extensive laboratory experience for the students. In these contexts, teachers typically offer some type of dissection experience, though faced with the three barriers described below.

1. Dissection not addressed in curricular materials and frameworks

A curricular gap exists. Though traditional and common within intermediate and secondary school biology classrooms, the practice of dissection is seldom mentioned within science education research, national curricular standards, and science frameworks. It has not had prominence in the past decade as a topic of importance. It does not appear in course outlines, and no major dialog concerning science curricula includes a consideration of dissection. Thus, there is no prominent platform where teachers and educational professors discuss methods for presenting laboratories involving dissections or alternatives. Teachers need to figure out for themselves how to structure these laboratories in their classrooms.

2. Phase-out of teaching resource centers

Instrumental and technical supportive resources for science laboratories have been sharply reduced across recent decades. County educational districts formerly provided resource materials that were integrated with specified laboratories for lesson plans and supported by specialists providing assistance with subject matter. These centers providing teaching resources have been dismantled. Teachers are on their own to acquire and accumulate teaching materials when needed to enhance their courses. The small budgets that are provided are only sufficient for purchasing a few clerical supplies. Abundant resources are available commercially, but they are costly and not presented as an integrated set of resources for high school biology (Weng et al., 2004a, 2004b). A few resources are available on loan, for example from Animalearn (2005) or the Humane Society of the United States (2005), but this requires planning well ahead and scheduling for particular lesson plans. The gap in the curricula and resources for science laboratories sets a stage for the third barrier.

3. Teachers' goal to motivate and interest students

To teachers, supplying motivating and informative materials

for students in classrooms is of prime importance. Teachers enter the profession dreaming of motivating students to learn, but are hampered in achieving their dream. They seek to inspire their students. In high school biology, a worthwhile laboratory exercise that would mark a quality experience for students is difficult for teachers to muster. Whether to use animal specimens and other resources in high school classrooms is not considered within the texts of curricular standards and science frameworks, nor are such resources and relevant expertise offered by school districts. Thus, the teachers' highest goal of inspiring their students in biology becomes ever more unattainable. A common question from teachers is, "How can we engage them?"

Solution to adoption of alternatives

The curricular requirements for teaching some laboratories in high school biology are relatively simple and straightforward, and could improve basic education in biology for students across the United States. In crafting a solution, it is critical to recognise that teachers have limited time available for laboratories on these topics; the resources required to meet their needs are not great. Even five outstanding laboratories produced in software and made freely available on the web could revolutionise biology laboratories in many classrooms. As a start, appealing software on the virtual mammal covering five basic laboratories on the skeletal-muscular, respiratory, digestive, nervous, and circulatory systems would provide a solid basis of education. With additional resources, the urinary, lymphatic and immune, skin, and endocrine systems could be added, plus organs such as the heart, brain, lungs, kidney, eyes, and ears, and joints such as the knee, hand, and foot, addressing the major needs of high school teachers. The teachers could complement these materials with other resources that they acquire.

Considering the number of students who could benefit from these improved teaching resources, the cost of preparing such software would be small. One possibility could be an increased commitment to this objective by animal advocacy groups (Fleischmann, 2003).

Conclusions

Teachers of high school biology often retain the traditional pattern of offering dissection of an animal. They seek new and improved resources, and consider using those that are easily accessible. But they are provided almost no budget for their laboratories, so are left to improvise and scavenge when designing the laboratories they offer. The disappearance of resource centers from school districts isolates teachers as they approach this quandary. The overriding motivation of teachers is to stimulate their students to enjoy learning. Dissection remains a favored avenue for providing an engaging experience to students. Production of five to ten outstanding software units on basic physiological systems could establish a solid foundation for secondary school biology laboratories everywhere.

References

Adkins, J. and Lock R. (1994). Using animals in secondary education – a pilot survey. *J. Biol. Educ. 28,* 48-52.

Akpan, J. P. and Andre, T. (2000). Using a computer simulation before dissection to help students learn anatomy. *J. Computers Math. Sci. Teaching 19*, 297-313.

Animalearn. Website accessed 08/15/2005: http://www.animalearn.org.

AVAR Alternatives in Education Database. Website accessed 08/15.2005: http://www.avar.org/.

Balcombe, J. (1997a). Student-teacher conflict regarding animal dissection. *Am. Biol. Teach. 59*, 22-25.

Balcombe, J. (1997b). Alternatives in education: overcoming barriers to acceptance. In L. F. M. van Zutphen and M. Balls (eds.), *Animal Alternatives, Welfare and Ethics* (441-444). Amsterdam: Elsevier Science.

Balcombe, J. (2000a). A global overview of law and policy concerning animal use in education. In M. Balls, A. M. van Zeller, and M. Halder (eds.), *Progress in the Reduction, Refinement and Replacement of Animal Experimentation* (1343-1350). Amsterdam: Elsevier.

Balcombe, J. (2000b). *The Use of Animals in Higher Education: Problems, Alternatives & Recommendations*. Washington, DC: Humane Society Press.

Balcombe, J. (2001). Dissection: the scientific case for alternatives. *J. Appl. Anim. Welf. Sci. 4*, 117-126.

Barr, G. and Herzog, H. (2000). Fetal pig: the high school dissection experience. *Soc. & Animals 8*, 53-69.

Bowd, A. D. (1993). Dissection as an instructional technique in secondary science: choice and alternatives. *Soc. & Animals 1*, 83-89.

Carolina Biological Supply Company. Website accessed 08/17/2005; http://www.carolina.com/.

DeBoer, G. E. (1991). *A History of Ideas in Science Education: Implications for Practice*. New York: Teachers College Press, Columbia University.

Fleischmann, K. R. (2003). Frog and cyberfrog are friends: dissection simulation and animal advocacy. *Soc.& Animals 11*, 123-143.

Gardiner, P. G. and Farragher, P. (1999). The quantity and quality of biology laboratory work in British Columbia high schools. *Sch. Sci. Math. 99*, 197-204.

Hart, L. A., Wood, M. W. and Weng, H.-Y. (2005). Mainstreaming alternatives in veterinary medical education: resource development and curricular reform. *J. Vet. Med. Educ.*, in press.

Humane Society of the United States. Humane education loan program. Website accessed 08/17/2005: http:www.hsus.org/animals_in_research/animals_in_education/humane_education_loan_program_help/.

InterNICHE. Alternatives loan system. Website accessed 08/17/2005: http://www.interniche.org/alt_loan.html.

Jukes, N. and Chiuia, M. (2003). *From Guinea Pig to Computer Mouse*. Leicester, England: InterNICHE.

Kinzie, M. B., Strauss, R. and Foss, J. (1993). The effects of an interactive dissection simulation on the performance and

achievement of high school biology students. *J. Res. Sci. Teaching 30*, 989-1000.
Klestinec, C. (2004). A history of anatomy theaters in sixteenth-century Padua. *J. Hist. Med. Allied Sci. 59*, 375-412.
Kline, A. D. (1995). We should allow dissection of animals. *J. Agric. Env. Ethics 8*, 190-197.
Langley, G. R. (1991). Animals in science education – ethics and alternatives. *J. Biol. Educ. 25*, 274-279.
Lock, R. (1994). Dissection as an instructional technique in secondary science: Choice and alternatives. Comment on A. D. *Bowd. Soc. & Animals 2*, 67-73.
Lock, R. and Millett, K. (1992). Using animals in education and research – student experience, knowledge and implications for teaching in the National Science Curriculum. *Sch. Sci. Rev. 74*, 115-123.
NASCO. NASCO Dissection Materials. Website accessed 08/17/2005: http://www.enasco.com/dissection/.
Orlans, F. B. (1991). Use of animals in education: policy and practice in the United States. *J. Biol. Educ. 25*, 27-32.
Orlans, F. B. (2000). An ethical rationale for why students should not be permitted to harm or kill animals. In M. Balls, A. M. van Zeller, and M. Halder (eds.), *Progress in the Reduction, Refinement and Replacement of Animal Experimentation* (1323-1331). Amsterdam: Elsevier.
Ralph, C. L. (1996). The illusion of the instructional biology laboratory. *Am. Biol. Teach. 58*, 142-146.
Sapontzis, S. F. (1995). We should not allow dissection of animals. *J. Agric. Env. Ethics 8*, 181-189.
Singer, S. R., Hilton, M. L. and Schweingruber, H. A. (eds.) (2005). America's Lab Report: Investigations in High School Science. National Academies Press. Website accessed 08/17/2005: http://www.nap.edu/catalog/11311.html.
Smith, A. (2005). NORINA : A Norwegian Inventory of Audiovisuals. Website accessed 08/15.2005: http://oslovet.veths.no/fag.aspx?fag=57&mnu=databases_1.
Smith, W. (1994). Use of animals and animal organs in schools: practice and attitudes of teachers. *J. Biol. Educ. 28*, 111-118.
Tansey, E. M. (1998). "The Queen has been dreadfully shocked": aspects of teaching experimental physiology using animals in Britain, 1876-1986. *Adv. Phys. Educ. 19*, S18-S33.
Weng, H.-Y., Wood, M. W. and Hart, L. A. (2004a). Alternative resources in education. Website accessed 08/15/2005: http://www.vetmed.ucdavis.edu/Animal_Alternatives/altsearch.htm.
Weng, H.-Y., Wood, M. W. and Hart, L. A. (2004b). Pre-college educators. Website accessed 08/15/2005: http://www.vetmed.ucdavis.edu/Animal_Alternatives/educator.htm.
Zasloff, R. L. and Hart, L. A. 1997. Adapting animal alternatives from veterinary medical education to precollege education. In L. F. M. van Zutphen and M. Balls (eds.), *Animal Alternatives, Welfare and Ethics* (445-447). Amsterdam: Elsevier.
Zirkel, J. B. and Zirkel, P. A. (1997). Technological alternatives to actual dissection in anatomy instruction: a review of the research. *Educational Technology 10(3)*, 52-56.

Acknowledgements
We appreciate the candid participation of the intermediate and secondary school teachers of the Sacramento Valley and UC Davis colleagues who candidly offered their experiences and needs, both during focus group meetings and in follow-up conversations. Partial financial support for these meetings was provided by the Bayer Corporation and Animalearn. Workshops and seminars for teachers sponsored by Schools and Colleges for Advancing the Teaching of Science (SCATS) over the past few years have been helpful. David Anderson provided some bibliographic assistance. Consultation with Robert Garmston provided the view of an educational administrator, and Sue Garmston gave her perspective from working in the California State Department of Education.

Correspondence to
Dr. Lynette Hart
UC Center for Animal Alternatives
School of Veterinary Medicine
University of California
Davis, CA 95616, USA
e-mail: lahart@ucdavis.edu

Should Live Animals be Used when Educating Future Biomedical Scientists?

Jann Hau
University of Copenhagen and Rigshospitalet, Copenhagen, Denmark

Summary
Animals are used for various purposes in education of scientists including: 1. Demonstration of biological principles, 2. Surgical education and training, and 3. Education and training of biomedical scientists for licensing.
This paper advocates a gradual replacement of animals for demonstration of biological mechanisms with non-sentient materials. It explains why animals remain necessary for training of surgeons, and why it is important for the welfare of laboratory animals that future scientists have passed a course demonstrating that they can handle and restrain animals, and perform routine procedures on anaesthetised animals before they can be granted a license to perform animal experiments.

Keywords: education, training, licensing

Introduction

The present Proceedings demonstrate that refinement is fortunately no longer the Cinderella of the Three Rs (Russell and Burch, 1959) as Hau and Carver wrote in 1994 – a term later adopted by van Zutphen (1998). There is now a general awareness of the importance of refinement in ethics committees and in the scientific community, as demonstrated by an increasing implementation of the Three Rs in biomedical research (Hagelin et al., 1999a and b, 2003; Hau et al., 2001; Carlsson et al., 2004).

The use of live animals for teaching and training purposes is controversial. The animals used obviously do not directly contribute to important new biomedical discoveries. Indirectly, however, they certainly do contribute, just as they contribute to the reduction of human morbidity and mortality associated with surgical treatment. To the uninformed public it may seem that this particular use of animals could appropriately be replaced by audiovisual material, new elegant interactive computer-based learning programmes and artificial models of animals and man. The welcome, rapid development in these new tools certainly contributes to a reduction in the number of animals needed for teaching and training, but they cannot completely replace the need for live animals. It is important, however, to be aware of the development of new alternatives and to implement these continuously, replacing live animals in teaching and training programmes, as and when these new devices and programmes become available.

Animal use in teaching and training

Live animals are presently used for different purposes in the education of physicians, veterinarians and scientists. These include the following:

1. Demonstration of physiological and pharmacological principles
   Undergraduate medical and veterinary programmes
   Undergraduate and post graduate science programmes
2. Surgical training
3. Training biomedical scientists for licensing

1. Demonstration of physiological and pharmacological principles

Many medical schools and veterinary schools have traditionally used anaesthetised animals in non-recovery demonstrations in practical classes to demonstrate principles and mechanisms of physiology and pharmacology, and many still do. Many new medical schools, however, have chosen not to use live animals in the curricula for undergraduate students and many old schools are revising their practicals, substituting animal-based practicals with, for instance, a combination of interactive computer programmes and extended practicals in which the students themselves are the guinea pigs. That there is indeed a plethora of new alternatives to live animals in this particular field is obvious to the reader of these Proceedings.

The need for live animals in the education of pre- and postgraduate science students is perhaps more difficult to substitute 100% with non-sentient material without compromising the quality of education. It seems likely, however, that looking more carefully at this level may also yield a reduction benefit. FELASA's policy statement on animals in teaching advocates replacement whenever possible, "The use of live animals in other teaching programmes, e.g. pharmacology and physiology for undergraduate students, to demonstrate fundamental biological interactions should be replaced with alternative methods as and when these become available" (http://www.FELASA.org). From an ethical point of view it is important to emphasise that vertebrates used in undergraduate curricula are generally anaes-

thetised prior to the practicals and euthanised at the end of the practical. Since the animals are unconscious and anaesthetised, they are not subjected to pain, distress or suffering.

2. Surgical training

Surgical education and training in new surgical procedures and techniques require the use of live animals, but it is probably possible to reduce the need for animals by a more extensive use of cadaver material and increasingly elaborate artificial models available on the market.

Training in microsurgery is traditionally dependent on live anaesthetised rats, and, although the initial phases of a training programme may well make use of artificial models, the final stages of the training require live animals – traditionally rats. This is well recognised and, in the UK for example, routine training of surgical techniques on animals is normally only allowed under supervision of operations or experiments, with a few exceptions, including training in microsurgery (Home Office, 1986). It is difficult to envisage that live animals can be completely replaced for this purpose in the foreseeable future.

The use of animals to train manual skills has been the subject of debate in Europe for several years, and in certain countries it is not permitted. In at least one major European country, animals may not be used in the teaching of laparoscopic techniques. Early mortality and morbidity figures were disappointing in this country, because the surgeons had not been trained on live animals. Consequently, a very large number of surgeons from this country have now been trained in minimal access surgical techniques abroad. This clearly demonstrates that animals remain vital for training of surgeons in minimal access surgical procedures.

Small numbers of animals – traditionally pigs – are used to train surgeons in war surgery, an activity which unfortunately continues to be highly relevant for humanitarian reasons. Fully anaesthetised pigs are shot at, and surgeons train how best to repair the damages assisted by veterinarians responsible for adequate anaesthesia of the animals. These practical exercises are non-recovery and are thus not really an ethical issue.

The use of animals in experimental surgery must thus be considered vital for human health, and there are certain indications that there is an increasing demand from surgeons for access to training facilities. This may be associated with the trend towards shorter working weeks for surgeons in various European countries, making it difficult for them to obtain the necessary routine and requirements for specialisation through their activities at the hospitals.

Rapid advancements in surgical techniques often require access to animals for research and development projects.

3. Training biomedical scientists for licensing

Training scientists in the most humane handling and procedural techniques. The utilisation of live animals in experimental biomedical research requires a certain level of knowledge and manual skills of the scientist who is the responsible license holder. This is well known and recognised among researchers and veterinarians because of its importance for upholding the quality of scientific results (Cohen, 1966) and safe-guarding the welfare of the animals used. The current European legislation requires individuals involved in research performed on animals to have adequate education (ETS 123, Council of Europe, 1986). The level of competence needed is usually not defined in laws and regulations. These decisions are left to those responsible for the animal experimentation on an institutional, regional or national level. It has long been argued that education in laboratory animal science should be strengthened and harmonised in Europe (Rozemond, 1991), and that FELASA should take the lead in this process. FELASA has done so and worked out guidelines for curricula, at four consecutive levels of competence, which should lead to adequate knowledge and skills for different categories of scientific and technical staff. In an increasing number of European countries the practice is now that the level of competence required for scientific staff that allows independent work with animals is the FELASA category C curriculum (Wilson et al., 1995). Thus, all new scientists, predominantly PhD students, have to complete this course and a written exam before they are allowed to work independently on animal experiments and before they can obtain a license to work with laboratory animals. FELASA's accreditation board for teaching and training programmes and courses is presently accrediting an increasing number of C-courses throughout Europe.

Practical training in correct handling, sampling and administration of substances must be considered essential components of the FELASA C course for biomedical scientists, who wish to obtain a license for animal experimentation. The FELASA statement on the use of animals in training courses reads as follows: "FELASA categories C …. courses must contain practical classes in which the participants are taught humane handling and restraint as well as modern and humane methods for blood sampling, administration of substances, anaesthesia and euthanasia."

Requiring that scientists, who wish to obtain a license to experiment on animals, must pass a course in which they are taught – among other important things – how to handle animals in the most humane way possible, how to restrain them so that the animals do not feel fear, and how to euthanise animals as well as how to perform the most common routine procedures (injections and blood sampling) seems to be an absolute minimal requirement. The overall goal is to help students approach their animals with greater insight and confidence than if they had to acquire the necessary knowledge and experience on their own, supervised by their scientific supervisor. The courses aim to give the students the basic knowledge and experience necessary to successfully continue acquisition of further practical skills during their experimental work. It is sometimes argued that new scientists could learn these attitudes and skills from experienced scientists, but there are more cons than pros of this old-fashioned system. It would be difficult to establish an assessment system where the practical skills of the new scientist were objectively assessed before he/she was found competent. On-the-job-training would also result in an uneven quality of teaching with no uniform standards, and this is not an area where it is acceptable to learn from one's own mistakes. A quality assurance system would also be difficult to operate without implemented accredited standards for practical training.

The numbers of animals used in training programmes is miniscule compared with research use and it must be considered essential from an animal welfare point of view that new scientists are taught how to restrain animals without causing fear and stress to the animal, as well as how to anaesthetise them. Once the animals are anaesthetised, the course participants can learn all relevant sampling techniques as well as administration of substances, after which the animal is killed and dissected for comparative anatomy. An important feature of training courses is the instruction and training in methods of humane killing. Some euthanasia techniques require more skill than others, and, in particular, physical methods like cervical dislocation should only be attempted by experienced personnel who have received sufficient training to allow them to use these methods confidently and effectively.

Discussion

Although it is often claimed that there is strong opposition to the use of animals in training programmes, this is not supported by looking at course evaluation forms. The author has taught postgraduate students for more than 25 years at universities in Denmark, the UK and Sweden, and the vast majority of students are extremely positive about the use of animals in training programmes. It goes without saying that it is important to show respect for the animals used and to demonstrate thoroughly how animals are picked up, restrained and anaesthetised before the students attempt this themselves (Hau, 1999). In addition, such courses should also foster a caring attitude to animals (van Zutphen, 1991), and teach appropriate and ethical treatment of animals (Ninomiya and Inomata, 1998). The importance of habituating animals to all staff including the scientists with the aim of lowering animal stress and increasing the pain threshold must also be emphasised, and it is important to teach how positive reinforcement training can help lower the stress of the animals as documented by Lambeth and co-workers (2005). Exercises with dummies, videos or interactive computer software cannot fully substitute for practice on live animals (Adam, 1993). Lack of proper training of future scientists may add variance in research results and increase the risk of future malpractice, thus compromising animal welfare. An introduction to the personal practice of the most efficient and humane experimental techniques through a high quality course ensures that future scientists will know how to treat their experimental animals with maximal consideration of animal welfare (Hagelin et al., 2000a).

The students attending our FELASA category C courses at Uppsala University 1997-2000 were asked to complete a detailed course evaluation in connection with the written exam at the end of the course (Carlsson et al., 2001). The forms were collected and processed anonymously. After completing the course, the majority of the students (>93%) were convinced that the subject was of great importance ($p<0.0001$). None of the students considered it to be of little importance after attending the course. It is noteworthy in the present context that more than half of the students (57%) found practical handling including experimental procedures the most important topic taught during the course. In particular, hands-on laboratory activities, like handling of and procedures (sampling and administration) on rats and mice scored highly (averaging >3.8 on a scale from 1 to 4). These results have been confirmed by a recent follow-up study (Abelson et al., 2005).

Compulsory courses are challenging, because they are not necessarily attended out of interest and motivation. The mandatory course in laboratory animal science taught at Uppsala University is no exception. It should be emphasised that since these courses are mandated, they have to be effective and perceived as highly relevant. There are few previous studies on attitudes towards courses in laboratory animal science. However, the Uppsala results seem to confirm the findings in a survey performed by the American Medical Association where 91% of physicians stated that the use of animals had been important for their training and 93% expressed support for continued use of animals in medical education (American Medical Association, 1989). Studies have shown that experienced teachers find that hands-on laboratory activities with animals add significantly to learning biology (Keiser et al., 1991; Mayer and Hinton, 1990; Offner, 1993). A recent study of veterinary students suggested that out of those who changed their view on whether or not it is morally acceptable to use animals for teaching, 26% became more receptive to the use of animals in research. The most frequent comments were that the course had given them a greater knowledge, a better understanding of the necessity of using animals for medical progress, more accepting and tolerant views and a less prejudiced view of the use of animals (Hagelin et al., 2000b).

References

Abelson, K. S. P., Hau, J. and Carlsson, H. E (2005). *Undergraduate and postgraduate students' responses o mandatory courses (FELASA category C) in laboratory animal science 1997-2003*. In Proceedings from the FELASA meeting in Nantes 2004, in press.

Adam, G. (1993). The role of education in understanding the use of animals in research. *Neuroscience 57*, 201-203.

American Medical Association (1989). *Survey of Physicians' Attitudes Toward the Use of Animals in Biomedical Research.* Chicago: American Medical Association.

Carlsson, H. E., Hagelin, J., Hoglund, A. U. and Hau, J. (2001). Undergraduate and postgraduate students' responses to mandatory courses (FELASA category C) in laboratory animal science. *Lab. Animals 35 (2)*, 188-193.

Carlsson, H. E., Hagelin, J. and Hau, J. (2004). Implementation of the Three Rs in biomedical research. *Vet. Record 154 (15)*, 467-70.

Cohen, B. J. (1966). Training Graduate Students in the use and care of Animals. *Federation Proceedings 25,* 1473-1476.

Council of Europe (1986). *European Convention for the Protection of Vertebrate Animals used for Experimental and other Scientific Purposes, ETS 123*. Strasbourg, Council of Europe.

Hagelin, J., Carlsson, H. E. and Hau, J. (1999a). Increased efficiency in use of laboratory animals. *Lancet 353 (9159)*, 1191-1192 APR 3.

Hagelin, J., Carlsson, H. E. and Hau, J. (1999b). Dutch university scientists increase research productivity using fewer animals per project. *Neth. J. Med. 55 (5)*, 212-214.

Hagelin, J., Carlsson, H. E. and Hau, J. (2000a). The importance of student training in experimental procedures on animals in education. *Scandinavian Journal of Laboratory Animal Science 27*, 35-42.

Hagelin, J., Hau, J. and Carlsson, H. E. (2000b). Attitude of Swedish veterinary and medical students to animal experimentation. *Vet. Record 146*, 757-760.

Hagelin, J., Carlsson, H. E. and Hau, J., (2003). Ethics' committees' refining influence on animal experimentation in Sweden. *Lab. Animals 37,* 10-18.

Hau, J. and Carver, J. F. A. (1994). Refinement in Laboratory Animal Science: Is it a Cinderella subject, and is there conflict and imbalance within the 3Rs? *Scandinavian J. Laboratory Animal Science. 21*, 161-167.

Hau, J. (1999). Humane endpoints and the importance of training. In C. F. M. Hendriksen and D. B. Morton (eds.), *Humane Endpoints in Animal Experiments for Biomedical Research (71-74)*. London: RSM Press.

Hau, J., Carlsson, J. E. and Hagelin, J. (2001). Animal research – Ethics committees have influenced animal experiments in Sweden. *Br. Med. J. 322 (7302)*, 1604-1604.

Home Office (1986). Guidance on the Operation of the Animals (Scientific Procedures) Act 1986. Presented to Parliament by the Secretary of State for the Home Department pursuant to Act Eliz. II 1986 C.14 Section 21 (Animals (Scientific Procedures) Act 1986), Appendix F, Education and training of personnel under the Animals (Scientific Procedures) Act 1986. (Available at: http://www.official-documents.co.uk/document/hoc/321/321-xf.htm).

Keiser, T. D. and Hamm, R. W. (1991). Forum: Dissection: The Case For. *Science Teacher 58*, 13-15.

Lambeth, S., Hau, J., Perlman, J. E. et al. (2005). Positive Reinforcement Training Affects Hematologic and Serum Chemistry Values in Captive Chimpanzees (Pan troglodytes). *Am. J. Primatol.*, in press.

Mayer, V. I. and Hinton, N. K. (1990). Animals in the classroom: considering the options. *Science Teacher 57*, 27-30.

Ninomiya, H. and Inomata, T. (1998). Current uses of laboratory animals in Japan and alternative methods in research, testing and education. *Applied Animal Behaviour Science 59*, 219-225.

Offner, S. (1993). The importance of dissection in biology teaching. *American Biology Teacher 55*, 147-149.

Rozemond, H. (1991). Veterinarians and the use of laboratory animals. *Der Praktische Tierarzt 72*, 892-894.

Russell, W. M. S. and Burch, R. L. (1959). *The principles of humane experimental technique*. Special Edition (1992). Publrs. UFAW.

Wilson, M. S., Berge, E., Maess, J. et al. (1995). FELASA-recommendations on the education and training of persons working with laboratory animals: category A and category C. Reports of the Federation of European Laboratory Animal Science Associations Working Group on Education accepted by the FELASA Board of Management. *Laboratory Animals 29*, 121-131.

van Zutphen, B. F. M. (1991). Education and Animal Experimentation. In C. F. M. Hendriksen and H. B. W. M. Koëter (eds.), *Replacement, Reduction and Refinement: Present Possibilities and Future Prospects* (119-125). Amsterdam: Elsevier.

van Zutphen, L. F. M. (1998). Refinement, the Cinderella of the 3Rs. In P. N. O. Donoghue (ed.), *The Ethics of Animal Experimentation*. London: EBRA.

Correspondence to

Jann Hau, Professor Dr. Med.
Head of Department of Experimental Medicine
University of Copenhagen and Rigshospitalet
3 Blegdamsvej
2200 N Copenhagen
Denmark
e-mail: jhau@emed.ku.dk

Session 1.2
Replacement alternatives in education: Animal-free teaching

Alternatives to Animal Experimentation in Undergraduate Curricula at Medical Schools – Analysis of Current Trends in the Czech Republic

Miroslav Cervinka[1], Emil Rudolf[1] and Zuzana Cervinkova[2]
[1]Charles University Faculty of Medicine, Department of Medical Biology and Genetics, Hradec Kralove, Czech Republic;
[2]Charles University Faculty of Medicine, Department of Physiology, Hradec Kralove, Czech Republic

Summary
Undergraduate medical students should be both theoretically and practically informed about the existence of alternatives to the use of animals in research and in education. Therefore we have prepared a course based on the 3Rs concept. In this course students learned and practically mastered the following topics:
- *The 3Rs concept – scientific background, ethical and legislative considerations*
- *Mammalian cells cultivated in vitro as an alternative to experiments on animals*
- *Non-invasive student self-experimentation as an alternative*
- *Screen-based alternatives (interactive computer programmes) as an alternative*
- *Proper use of laboratory animals*

We prepared a written anonymous questionnaire to evaluate students' opinions on the course and their attitudes towards the alternatives. The results of the survey showed that our students were generally satisfied with our course and it seems that both experiments with cells in vitro and human experimentation could be suitable alternatives in medical education.
Due to the fact that we organised similar surveys several times during the last 15 years, it was possible to analyse changes in the students' attitudes during that period. The results showed one obvious general tendency: current less strict opposition of students towards animal experimentation, with a substantial part of our respondents even requiring animal experimentation. Reasons behind these changes are discussed.

Keywords: alternatives, animal experiments, education, faculty of medicine, curricula

Introduction

Animals used for teaching and education purposes represent only about 1% of all laboratory animals used for any purpose. Despite this fact, teaching alternatives to the use of laboratory animals is very important, because it forms the attitudes of students for their whole life (van der Valk, 1999). The main goal of this article is to present a brief historical survey of changes in the use of laboratory animals in education at the Faculty of Medicine in Hradec Kralove over the last 15 years. Subsequently, the current situation at the faculty is described, including positive achievements in this field as well as speculations about future trends. Focus will mainly be on the situation at our faculty, considering that the situation is very similar in all

seven medical faculties in the Czech Republic. We believe that our data are relevant for all undergraduate students of medicine in our country.

Situation in the 1990's

Use of animals in education at universities has a longstanding tradition in our country; it used to be a standard element of the curriculum at all medical faculties. Before 1990, laboratory animals were easily obtained and there was no legislation concerning their use. Generally, in this field there were no regulations and practically no restrictions.

After the revolution in November 1989, the attitude of our society changed substantially and the protection of animals became one of the important though controversial issues. Our federal parliament voted for a new Animal Protection Law in April 1992 (Czech National Council, 1992). The law was in very good concordance with EC directive 609-1986 (Council of Europe, 1986) and has since been amended several times to be in full harmony with current EU legislation.

At the Faculty of Medicine in Hradec Kralove we tried to promote changes in the attitude of teachers and students towards alternatives immediately after the revolution. At that time we had personal contacts with leaders in the field of alternatives, namely with Nottingham University Medical School (Department of Human Morphology), FRAME, Free University of Berlin and ZEBET. Together, we prepared a proposal for the TEMPUS Joint European Project No. 1485 "Alternatives to Experiments with Animals in Medical Education". This project affected the situation at our faculty substantially.

The main goal of this Joint European Project was to introduce the 3Rs concept and practical approaches to alternatives to animal experiments in the education system, and promote:

- Improvements in the curricula at the medical faculty
- Work of committees for proper animal use at the medical faculty
- Training in the practical use of alternative methods
- Production and exchange of new teaching materials concerning alternatives
- Improvements in husbandry and care for laboratory animals
- Obligatory regulation of experimentation with animals

The results of this project were summarised at the TEMPUS evaluation meeting in Hradec Kralove in May 1994 and published as a book (Cervinka and Balls, 1995). We are obliged to TEMPUS Office and to our friends in EU countries (M. Balls, H. Spielmann, R. Clothier, M. Liebsch, and others) for their longstanding support.

As a result of these and other activities, there was a very marked decrease in the use of laboratory animals for teaching purposes at our faculty. In table 1, the numbers of laboratory animals used for teaching purposes in the period of 1991-2004 are given. It is clear that there was a very rapid decline in the use of laboratory animals at the beginning of the 90's, from 1830 to 160 animals per year. Since that time the numbers have remained almost constant. In 1991, animals were used for teaching at eight departments in 36 practical lessons. Since 1995, laboratory animals were used for teaching purposes at only two departments – the Department of Physiology and the Department of Pharmacology.

Situation in 2000

At the turning point of the millennium, the situation at our faculty was stable. The system of proper animal use continues without disturbances. All students and teachers at our faculty follow strict regulations. Basic characteristics of our curricula can be summarised as follows:

- No use of dead animals for dissections
- No use of animals for surgical training
- Very limited harmful use of living animals. The Animal Welfare Committee must approve each proposal for the use of animals in education every year.
- Students may choose an animal-free alternative to these practical classes.

After many years of teaching surgery and other clinical subjects without using animals, it is obvious that the level of practical skills of our graduates remains the same. With this we can be satisfied. Nevertheless, we realise that most regulations at our faculty are restrictive. We think that restrictions are not the best way to change the attitude of people, and, in particular students. Therefore, we try to offer a positive alternative.

Current situation

We believe that undergraduate medical students should be both theoretically and practically informed about the existence of alternatives to the use of animals in research and in education. Students are exposed to a clear and thorough consideration of the advantages and disadvantages of alternatives. Each student has practical experience with alternative *in vitro* methods. Simultaneously, they should absorb basic facts about experiments on animals. Therefore, we have prepared a course module based on the 3Rs concept of alternatives. It is a joint co-operative action of two departments, one using animals in teaching (Department of Physiology) and the other using alternative *in vitro* methods in teaching (Department of Biology).

Tab. 1: Numbers of laboratory animals used for teaching purposes in the period of 1991-2004 at Charles University Faculty of Medicine in Hradec Kralove.

	1991	1993	1995	2000	2004
Rat	470	85	20	20	20
Mouse	740	490	120	30	0
Guinea pig	100	0	0	0	0
Rabbit	380	124	20	10	17
Frog	110	40	0	0	0
Dog	32	0	0	0	0
Total	1832	739	160	60	37

The Alternatives To Animals Module (ATA Module) is obligatory for all students of general medicine in the second study year. It consists of 3 practical classes and one seminar in physiology (together 12 teaching hours) and 2 practical classes in biology (6 hours). Each academic year about 150 students pass this course module. During the teaching activities we facilitate discussions among students on the topic of alternatives. Our ideas on this module were already presented at the Workshop on Teaching Alternatives in Warsaw in 2002 and at the 4th World Congress on Alternatives in New Orleans in 2002 (Cervinka and Cervinkova, 2003).

In the ATA Module students learn and practically master the following topics in five teaching blocks:

1. Concept of Alternatives. All students are informed about our goal – to promote progress towards the refinement, reduction and replacement of animal experiments. Extra time is spent on explanation of the 3Rs principle, on scientific background, ethical and legislative considerations, and on information about alternatives on the internet.

2. Proper use of animals. Currently, animals are used only in the following practical lessons: "Nervous and humoural regulation of the blood pressure in the rabbit" and "Perfusion of the rat heart *in vitro* – effects of drugs". These practical lessons are also used to deliver information about proper housing and handling of laboratory animals, about humane practice in conducting animal experiments and about legal and ethical aspects of animal experiments. Students can choose between animal and non-animal alternatives during practical lessons, with non-animal alternatives having the same educational level.

3. Computer models as an alternative in teaching (theory of systems, advantages of modelling, practical use of computer models as an alternative to animal experimentation). To carry out this module it was essential to build a new computer laboratory with new computer teaching programmes and to establish a new student teaching laboratory for cultivation of cells *in vitro*. For both of these activities, we received financial support from the Czech ministry of education. Also, we very much appreciate help from InterNICHE, Nick Jukes and other co-operators who helped with know-how and the selection of suitable computer programmes. The following programmes are now available: computer simulation of action potential (AXONLAB), modelling of neuromuscular transmission (CONDUCT), computer modelling of blood pressure regulation, computer simulation of oxygen transport, and computer simulation of kidney functions.

4. Non-invasive students' self-experimentation as an alternative method to animal-based models. Practical classes performed on student volunteers are the most common type of practicals in physiology, including haematology, neurology, physical examination of the cardiovascular system, physiology of the senses, oral glucose tolerance test and measurement of basal metabolic rate.

5. Practical use of mammalian cells cultured *in vitro* as an alternative in research and testing. In this part of the ATA module students are educated in state-of-the-art *in vitro* techniques. During the practical course organised at the Department of Biology, we introduce students to *in vitro* methodology and demonstrate the use of cultivated cells in various applications, such as toxicity assessment. Students are invited to discuss advantages and disadvantages of various models employed in the field of toxicology, especially with regard to recent developments in respective EU policies.

Special emphasis is placed on evidence-based *in vitro* models and on validation of alternative methods. For these purposes students are acquainted with the main information sources on the internet – each student individually searches for new information about alternatives. After this practical class each student knows where to find recent data about alternatives (Altweb, Norina).

Training in the practical use of alternative methods is organised at the Department of Medical Biology and Genetics and starts with an introductory lecture on *in vitro* cultivation of mammalian cells and their application in teaching and research, followed by an excursion to the *in vitro* toxicology laboratory. Finally, in one practical lesson students themselves perform testing of cytotoxicity of a selected chemical on the Hep2 cell line *in vitro*. Examples of these experiments include the photodynamic effect of selected chemicals on human and murine cell cultures and assessment of toxicity of some materials/substances used in human health care. These simple experiments serve to demonstrate usefulness of *in vitro* cell cultures in numerous areas of scientific experimentation. In our opinion, this hands-on experience greatly enhances the students' understanding of alternatives.

Student's evaluation

Since all our efforts of teaching alternatives are in a way unique in the Czech Republic, we need feedback from our students. This year we prepared a written anonymous questionnaire to evaluate students' opinions on the course and their attitudes towards alternatives. Due to the fact that we organised similar surveys in 1995 and 2000, it was possible to analyse changes in the students' attitudes over the past 15 year period.

The results of the survey showed that our students were generally satisfied with our new module, and it seems that both experiments with cells *in vitro* and human experimentation could be suitable alternatives in medical education. Generally, it is very positive that more than 107 students of general medicine participated in our enquiry, i.e. 70% of all students in the second academic year. Another positive point is the fact that very few students selected the option "I cannot answer" in any of the questions. This proves that all questions were precisely formulated and it is a strong indication that our students have firmly fixed opinions on the use of laboratory animals. The questions were grouped in four clusters.

The first cluster of questions was oriented on the use of laboratory animals in general. The majority of respondents (58%) believed that the use of laboratory animals for research purposes is essential (indispensable). On the other hand, only 44% of respondents believed that laboratory animals are essential in education, and 53% of respondents did not agree with this statement. Complete elimination of laboratory animals from the education process is supported by 32% of respondents, with 71% of respondents opposing this.

The second cluster of questions was focused on the use of laboratory animals at our medical faculty. It is heartening that the majority of students (72%) consider the extent of use of laboratory animals at our faculty adequate. 59% of respondents do not agree with the statement "The use of laboratory animals should be reduced" and 84% of respondents do not agree with the statement "The use of laboratory animals could be increased". Almost all students (96%) appreciate the option to refuse to work with animals without any sanctions.

The third cluster of questions focused on alternatives to laboratory animals. The majority (76%) of students voiced the opinion that they are sufficiently informed about alternatives to laboratory animals. Virtually all students (with only one exception) personally agreed with and accept the 3Rs principle. Practically all students considered incorporation of teaching about alternatives as important and essential in the curriculum at faculties of medicine.

In comparison with previous years, one general tendency is obvious, students are currently less strict in their opposition against animal experimentation, and a substantial part of our respondents even required animal experimentation.

Students were slightly more critical in the assessment of our lectures, seminars and practicals about alternatives. Still, the majority (78%) of students agreed with the general scope and focus of the course and with its syllabus. One special question was aimed at the only practical class where students work with living animals. This practical class is "Direct measurement of blood pressure in the rabbit". At the end of this class, animals under anaesthesia are sacrificed. Sacrifice notwithstanding, 71% of respondents considered this class appropriate. 83% of respondents very positively acknowledged that after asking, they were presented with special videotaped instructions on the experiment. The second part of the teaching block – practical training in the use of alternatives (i.e. *in vitro* cytotoxicity assessment on cells cultured *in vitro*) was positively accepted by 75% of respondents.

When all these responses are taken into account, it is clear that students themselves consider teaching of alternatives to laboratory animals at our faculty as fully acceptable. Students were asked to use the same grading system as in high school to evaluate our teaching module on alternatives; 88% of students labelled this module as excellent or very good.

In the next section of the questionnaire we tested our students' knowledge on alternatives. Virtually all students were able to describe the exact meaning of the 3Rs. 68% of respondents knew that experiments on animals should be probed by the national committee for the use of animals and not by the head of department or the dean. When asked to enumerate types of alternatives to laboratory animals in education, the students displayed less knowledge. The vast majority of students correctly named computer models and cells cultured *in vitro* as examples of alternative teaching models, but only 23% of respondents correctly noted the third possibility – non-invasive experiments on volunteers.

In the last part of the questionnaire students could add personal comments about the use of laboratory animals at our faculty or put down some suggestions for improving the situation. This opportunity was used by only 50% of respondents, and their opinions differed greatly. One explanation for this may be that comments and suggestions came from students with well-defined standpoints only. Generally, students recommended small working groups, extended time for practicals, more animals, or, alternatively, complete replacement of animals. The main advantages mentioned in connection with the use of laboratory animals were: visualisation, first contact with living animals, elimination of the fear of "cutting a living creature", authentic experience, etc. Only in a few cases did the students acknowledge the opportunity of being informed about scientific methodology.

Being teachers, we can be very pleased with the fact that students appreciate it when teachers try to minimise the suffering of animals, work very carefully with animals and organise practical work well.

Few students stated that they had had a negative experience – not properly effected anaesthesia or insufficient experience of some assistant teachers. There were clearly differences between students of different study groups who had different teachers. To rectify this situation, we will have to pay more attention to harmonisation of skills and attitudes of individual teachers.

Conclusions – future trends

Nowadays, the situation at our faculty is stabilised and the use of laboratory animals in education is very limited (about 20 rats per year). We have implemented international standards in husbandry and care for laboratory animals and the obligatory regulations for experimentation with animals are accepted. Non-animal alternatives are available and students can choose between animal and non-animal alternatives during their practical lessons. All students have to pass the obligatory teaching module about alternatives to animals and receive balanced information on the proper use of laboratory animals and advantages and disadvantages of alternatives in a harmonised teaching module.

In the future, we will emphasise not only ethical issues but also scientific merits. We believe, that there are many good scientific reasons to replace some animal experiments with some *in vitro* methods. However, the same point of view will dictate that some animal experiments, which are indispensable from the scientific standpoint, will remain. In light of our current experiences, one question remains open: Is animal-free teaching at medical faculties the ultimate goal, or are very limited experiences with living animals essential for students of medicine?

References

Cervinka, M. (2004). Situation in the Czech Republic: An Example of a Country in Transition. *ATLA 32, Suppl. 1*, 595-597.

Cervinka, M. and Balls, M. (1995). *Alternatives to animal experimentation*, 99 p. 1st ed. Hradec Kralove: Nucleus.

Cervinka, M. and Cervinkova, Z. (2003). Alternatives to animal experimentation in undergraduate curricula at medical schools. *ALTEX 3*, 164-165.

Cervinka, M. and Cervinkova, Z. (2003). *Alternatives to Animal Experimentation in Undergraduate Curricula at Medical*

Schools. In: Workshop on Alternatives to the Use of Animals in Higher Education – Programme and Abstracts. Warsaw, Poland, October 18-20, 23.

Council of Europe (1986). *European Convention for the Protection of Vertebrate Animals Used for Experimental and Other Scientific Purposes*. Strasbourg, France: Council of Europe.

Czech National Council (1992). Act No. 246/1992 S.B "On the protection of animals against abuse".

Valk van der, J. et al. (1999). Alternatives to the Use of Animals in Higher Education. The Report and Recommendations of ECVAM Workshop 33. *ATLA 27*, 39-52.

Acknowledgement

This work was supported by the Czech Republic Ministry of Education Research Project MSM 0021620820.

Correspondence to

Prof. M. Cervinka
Department of Medical Biology and Genetics
Charles University Faculty of Medicine
Simkova 870
50038 Hradec Kralove, Czech Republic
e-mail: cervinka@lfhk.cuni.cz

University of Virginia Medical School Replaces Canine Lab with Human Patient Simulator

Rooshin B. Dalal[1], Megha Shah Even[2], Chad B. Sandusky[2] and Neal D. Barnard[2]
[1]University of Virginia School of Medicine, Charlottesville, Virginia, USA; [2]Physicians Committee for Responsible Medicine, Washington, DC, USA

Summary
Until recently, the University of Virginia School of Medicine (UVa) taught emergency surgical procedures using a canine laboratory. Medical students, working with faculty, administration, and community members, eliminated the use of canines and implemented a new life-saving techniques course using a human patient simulator and other stand-alone stations, allowing students to practice repeatedly without using animals. Replacement of this canine lab marked a turning point for medical education at UVa, and follows a general trend since 1994 of declining use of animals in medical education in the U.S.

Keywords: replacement, medical education, live animal labs, surgical training, human patient simulators

Introduction

Live animal laboratories were traditionally used for basic skills training in physiology, pharmacology, and/or surgery in the medical school curriculum. In 1985, virtually all 126 medical schools in the U.S. offered a live animal laboratory as a requirement for physicians-in-training. Medical students, depending on the exercise, injected drugs such as epinephrine, into an anesthetised animal and monitored physiological responses such as heart rate and vascular tone. Medical students also practiced surgical skills, such as suturing, IV placement, and chest tube insertion on anesthetised animals. Canines were typically used and the animals were euthanised at the end of the laboratory. Between 1985 and 1994, a decline in animal use in medical education was seen (Ammons, 1995; Wolfe et al., 1996). A continuation of this trend was demonstrated in a 2001 survey in which 68% of U.S. medical schools had eliminated live animal laboratories for student training. Furthermore, when animal laboratories were offered, participation was optional in most cases (Hansen and Boss, 2002). Recent surveys conducted by the Physicians Committee for Responsible Medicine (PCRM) indicate that over 80% of U.S. medical schools have eliminated all live animal laboratories from their curricula (unpublished data). The remaining 20%, although not requiring participation, still offer live animal labs to civilian medical students. The dramatic decline in animal use over the past 15 years has been attributed to the cost of using animals, curricular expansion, and ethical concerns (Hansen and Boss, 2002).

The University of Virginia (UVa) School of Medicine in Charlottesville offered a canine laboratory to 3rd year medical students for training in emergency surgical skills. Our approach was to promote one R, *replacement*, of the 3Rs concept of *replacement*, *reduction* and *refinement* of animal use in education. We worked with the community to eliminate the outdated canine labs from the undergraduate medical curriculum at UVa and to replace them with superior, non-animal training methods.

Alternatives to live animal laboratories in medical education

Animal use in education is undoubtedly an ethical issue, and guidelines set forth by U.S. Institutional Animal Care and Use Committees state that alternatives to the use of animals in education are to be considered and used when available (National Research Council, 1996). Over the past decade, advances in computer software and simulation technologies have fostered the development of numerous, high-quality alternatives to animal laboratories for medical training. In addition to videos and ethically-sourced cadavers, virtual reality programmes and simulation models provide training that is humane, clinically-relevant, and more realistic than animal laboratories. The advantages of using human-based training methods include anatomical accuracy, the ability to achieve proficiency through repetition of procedures, a realistic clinical environment (Gordon and Pawlowski, 2002), and long-term cost savings.

Alternatives for basic skills training in physiology, pharmacology, and surgical skills are numerous and widespread. Medical physiology can be taught using interactive computer programmes in which students are able to monitor changes in physiological parameters, such as blood pressure and heart rhythms after administration of "virtual" drugs. One study compared student test scores for a physiology laboratory when using videodiscs or when using live animals and found no difference in scores between the two training methods (Fawyer et al., 1990). Another group compared computer-based training to animal laboratories for cardiovascular physiology and showed that students favoured computer-based training (Samsel et al., 1994).

These studies corroborate data from a survey of U.S. medical schools which shows a 50% decline in the use of live animals in medical school physiology and pharmacology from 1994 to 2001 (Hansen and Boss, 2002). Although the decline in live animal use in medical education overall is significant, the use of live animals in surgery courses was stable from 1987 to 2001 (Hansen and Boss, 2002). Recent surveys conducted by PCRM indicate that animal use for surgical training is declining and a number of studies demonstrating improved student performance of surgical skills using alternatives versus live animals have been conducted (McCarthy et al., 2002; Tsang et al., 1994). Moreover, in countries such as the UK, live animal use for the practice of surgical techniques has never been a tradition (Jukes, 2004). Instead, human anatomical simulators and/or perfused organ systems allow students to practice surgical techniques on anatomically accurate models with simulated blood flow. In addition, student self-assessment and clinical practicals are useful adjuncts that replace animal laboratories.

Chronology of events to replace the canine laboratory at UVa

Upon learning that the majority of US medical schools had eliminated terminal live animal laboratories from their curricula, medical students at the University of Virginia sought to replace their canine laboratory with humane, non-animal methods as well. The University of Virginia offered a "Life-Saving Techniques Laboratory" course to 3rd year medical students, in which emergency surgical skills were practiced on 96 beagles per year. The students were shown each of the following procedures once: chest tube insertion, venous cut-down for intravenous line placement, cricothyroidotomy, and splenectomy. Although the laboratory course was optional, the instructor described the laboratory as "one of the best experiences students will have in medical school." Conscientious objectors to the use of live animals for this purpose were given an opportunity to opt out of the lab only hours before the start of the lab. The procedure used by the course directors to identify such students was to have them raise their hands in front of the entire class. This intimidating environment, and the lack of opportunities to use alternatives, may have caused some medical students at UVa to fear academic consequences or embarrassment for not participating in the animal laboratory.

Efforts to replace the canine laboratory were initiated in October 2002 by a veterinary technician responsible for the care of the beagles. Comments from a few medical students regarding their unpleasant experiences with the canine laboratory prompted the veterinary technician to send an e-mail to first and second year UVa medical students about alternatives to the use of live animal laboratories. A canine laboratory was scheduled for January 2003 and third-year medical students were informed of the lab only 2 weeks prior to the start. Two medical students organised a presentation for January 31, 2003, entitled "Alternatives to Live Animal Use in Medical Education", given by Neal Barnard of PCRM. One hundred and fifty people, mostly medical students, residents, and faculty, attended the presentation. Lengthy discussions about the scientific and ethical concerns surrounding live animal labs followed. The University newspaper, *The Cavalier Daily*, published a front page story about student opposition to the canine laboratory, as well as letters to the editor, both in support of and in opposition to the animal lab.

The canine laboratories continued until August 2003, when two medical students presented the instructor of the "Life-Saving Techniques Laboratory" course with a list of alternatives for surgical skills training. Third year medical students in the surgical clerkship were asked their preference for a human cadaver lab or a canine laboratory; 50% of the students (n=18) indicated a preference for a human cadaver lab, 25% (n=9) preferred a canine laboratory, and 25% (n=9) had no preference either way. During the next session of the course, the instructor made arrangements for only 4 students to learn the procedures on a single human cadaver. No human cadaver alternative was offered after that initial lab, even though the 4 students that completed the lab had overwhelmingly favourable comments about their experience.

The continued failure of faculty and administration to permanently implement alternative training methods at UVa prompted medical students to organise another presentation on alternatives to animal use. In October 2003, a trauma surgeon presented various training methods for emergency surgical procedures that do not rely on animal use. This talk was countered two weeks later with a presentation by the course director justifying the use of canines for trauma training. Medical students and concerned citizens in the community formed an advocacy group called Citizens for Humane Medicine in November 2003. They began contacting the media to increase community awareness of this lab in an effort to replace it with more modern and effective training methods.

In January 2004, the local newspaper, *The Daily Progress*, printed a front page story describing the efforts of Citizens for Humane Medicine (CHM) to eliminate the canine laboratory. Media coverage of CHM's efforts continued for several weeks, followed by phone calls from community members supporting the replacement of the canine labs. In February 2004, the Dean of the Medical School announced a suspension of the canine labs pending a review by a newly formed committee. The 12-person committee, comprised of faculty, residents, and one medical student, voted to permanently eliminate canine laboratories in March 2004. The committee also recommended that alternative training methods be implemented similar to those being used at most other U.S. medical schools.

Results

A new, full-day course entitled "Life-Saving Techniques Workshop" was implemented by the Emergency Medicine Department in November 2004 to replace the canine lab that was previously conducted by the Surgery Department. The new course uses six workstations to teach emergency procedures. These include chest tube insertion, adult and paediatric airway intubation, surgical cricothyroidotomy, and vascular access in

various anatomical locations. Lastly, a human anatomical simulator called Human Patient Simulator™ (HPS; METI, USA) was purchased to give medical students a realistic environment in which to identify life-threatening situations and perform life-saving procedures (Gordon, 2000).

Conclusion

Replacement of canine surgical laboratories at UVa follows a general trend of declining animal use in U.S. medical education over the past 20 years. Efforts to eliminate crude animal laboratories for educational purposes at UVa took approximately two years. The involvement of dedicated medical students and community members was critical to its success. The surgical faculty at UVa were not amenable to the implementation of non-animal training methods, in spite of the availability of suitable alternatives and their use by over 80% of medical schools in the U.S. Ultimately, public pressure through sustained media coverage forced UVa to re-evaluate its curriculum. The result was to replace its canine laboratory with a superior, non-animal training method. We believe the strategies employed at UVa can be used as a guide for replacing animal laboratories at other medical schools and in other educational settings to achieve the goals of the 3R's concept of animal use.

References

Ammons, S. W. (1995). Use of live animals in the curricula of U.S. medical schools in 1994. *Acad. Med. 70*, 739-743.

Fawyer, A. L., Branch, C. E., Trentham, L. et al. (1990) A comparison of interactive videodisc instruction with live animal laboratories. *Am. J. Physiol. 259*, 511-514.

Gordon, J. A. (2000). The Human Patient Simulator: Acceptance and efficacy as a teaching tool for students. *Acad. Med. 75*, 522.

Gordon, J. A. and Pawlowski, J. (2002). Education on-demand: The development of a simulator-based medical education service. *Acad. Med. 77*, 751.

Hansen, L. A and Boss, G. R. (2002) Use of live animals in the curricula of U.S. medical schools: Survey results from 2001. *Acad. Med. 77*, 1147-1149.

Jukes, N. (2004). Replacement of harmful animal use in life science education: The approach and activities of InterNICHE. *ATLA 32 (Supplement 1)*, 511-515.

McCarthy, M. C., Ranzinger, M. R., Nolan, D. J. et al. (2002). Accuracy of cricothyroidotomy performed in canine and human cadaver models during surgical skills training. *J. Am. Coll. Surg. 195*, 627-629.

National Research Council (1996). *Guide for the care and use of laboratory animals.* Washington DC: National Academy Press.

Samsel, R. W., Schmidt, G. A., Hall, J.B. et al. (1994) Cardiovascular physiology teaching: computer simulations versus animal demonstrations. *Am. J. Physiol. (Advances in Physiology Education)*, 536-546.

Tsang, S. M., Caluda, M. J., Steinberg, S. M. et al. (1994). Laparascopic cholecystectomy: What's so special? *South. Med. J. 87*, 1076-1082.

Wolfe, M., Barnard, N. D. and McCaffrey, S. M. (1996) Animal laboratory exercises in medical school curricula. *ATLA 24*, 953-956.

Acknowledgements

The authors wish to thank the Citizens for Humane Medicine, dedicated medical students, and the animal care veterinary technician at UVa. We are also greatly appreciative of the support we received from Claudio Pinto at *The Daily Progress* and Simon Chaitowitz at PCRM.

Correspondence to

Rooshin B. Dalal, MD/PhD Student
University of Virginia School of Medicine
1300 Jefferson Park Avenue
Jordan Hall 3226
Charlottesville, VA 22908, USA
e-mail: rooshin@virginia.edu

RECAL: Creating Computer-assisted Alternatives Using a Sustainable Learning Objects Approach

David Dewhurst, Stewart Cromar and Rachel Ellaway
College of Medicine & Veterinary Medicine, University of Edinburgh, UK

Summary
The mainstay alternatives to using animals in higher education are multimedia computer-assisted learning (CAL) programmes simulating practical pharmacology classes. CAL development intrinsically ties the educational content and learning design to the authoring application. As technologies change, authoring applications become obsolete, leaving redevelopment at further expense as the only option.
RECAL is based on principles of standards, objects and reusability and is developing methods and tools to break this cycle by disaggregating existing CALs to separate their constituent learning objects from the runtime environment. This has improved their long-term viability and facilitated their adaptation by teachers to meet divergent learning needs.

Keywords: computer-assisted alternatives, pharmacology teaching, higher education

Introduction

Over the last twenty years the mainstay of technology-supported teaching and learning in the biomedical sciences has been interactive multimedia computer-assisted learning (CAL) programmes designed to support specific areas of the curriculum. In pharmacology and physiology a number of these programmes have been aimed at laboratory teaching of these subjects that typically use animals or animal tissues. In such cases the CAL programmes provide an alternative to the educational use of animals and are very much consistent with the 3Rs and replacement in particular. Although they are not the only form of alternative (see Gruber and Dewhurst, 2004 for a more detailed account), they are the major form of alternative available to teachers who use animals in teaching of pharmacology, probably the discipline that is the major user of animals in education. To be effective in reducing animal usage they must be acceptable to educators – computer-assisted alternatives must be shown to be educationally effective, they must meet the learning objectives of the laboratory session they were developed to replace, they must be able to be delivered to students in locations in which learning takes place and they must be low-cost. A number of studies have evaluated the effectiveness of e-alternatives. These suggest such resources are able to deliver low-cost, effective education that meets primary teaching and learning goals (Dewhurst, et al., 1994; Hughes, 2001; Fawver, et al., 1995; Guy and Frisby, 1992; Leathard and Dewhurst, 1995).

There is now a critical mass of high quality e-alternatives available to teachers (see www.eurca.org for a list of independently reviewed programmes) and many universities across the world have now substituted at least some animal experiments in their curricula. Most of the computer programmes currently available were developed in the 1990's and are marketed to universities in CD-ROM format. In most instances universities copy the programmes to their institutional server(s) and deliver the programmes to students across their local area network (LAN). With the now near ubiquity of the Internet and a shift to more (off-campus) student-centred learning activity in higher education, there is increasing demand to be able to access learning resources via the World Wide Web.

The fact that there is widespread availability of computer-based alternatives that have been demonstrated to be educationally effective and have received good independent reviews does not necessarily mean that teachers will actually use them and thereby reduce animal use. The usefulness of a CAL programme depends on the closeness of fit of the programme to the needs of the teacher and his or her willingness to adopt materials developed elsewhere. It is unlikely that teachers will find programmes that meet all of their needs. Anecdotally at least, many teachers want the opportunity to be able to tailor these programmes to meet their local context of use. Thus for instance, in a pharmacology simulation programme they might wish to add new or delete some of the test drugs, extend the dose range, add new tasks or delete current ones, display the data in a different format (e.g. extend the time base), translate the text into a language other than English, add formative assignments, etc. Similarly, small clips from videos of animal dissection or anatomy might be more useful and usable than the whole presentation. To date, the intrinsic constraints of the authoring tools used to create the alternatives make this editing process very difficult and expensive: editors would have to acquire the original programme's source code, purchase the appropriate authoring software, and they would need to find staff with the skills, time and experience to change the resource to meet their requirements. The redevelopment process would then be repeated at every university that needed a variation of the original CAL programme.

From this it is clear that today's educators require editable, web-based learning materials – properties which are not consistent with the LAN-based multimedia CD-ROMs of the last decade when programmes were developed using commercial

authoring programmes (such as Toolbook (Asymmetrix), Authorware (Macromedia) and Director (Macromedia)). The product was a compiled executable programme containing all of the learning and media assets (text, images, graphics, animations, video, audio, self-assessment questions), the sequencing of these assets, and the learning design. The educational content was therefore intrinsically linked to the technologies used to develop and deliver the programmes. Technology has changed rapidly (such as the move from DOS to Windows, from 16 bit to 32 bit processing and from VGA to XGA and above screen resolutions) and it continues to change. This ongoing flux means that while the educational content and learning design of a programme may still be valid, it too may be lost as its delivery mechanism becomes obsolete.

The RECAL project described here is developing methodologies that make use of new ways of abstracting and managing the educational content (assets), pedagogical design and run-time components of existing e-alternatives to allow for much greater longevity and flexibility of such materials. Here we present an introduction to the RECAL approach, its use of common standards and specifications to describe assets and educational activities and how it can 'future-proof' CAL materials and thus benefit both developers and users.

Methods: the RECAL process

The RECAL architecture

The RECAL process fuses new fast and light web-based multimedia technologies such as Macromedia Flash to develop the content with the use of XML and is based on common learning technology standards and specifications. Semantically-rich XML (extensible mark-up language), which can be saved as cross-platform text files and can be created and transported rapidly between applications, is used to encode and thereby preserve the content and design of CAL materials. This can, as a result, expand their lifespan significantly and enable more effective reuse.

The RECAL architecture is based on a number of disassociated components (the run-time engine (currently Flash), the text (XHTML files) and media assets which are stored on a web-server and the XML parameter files describing the sequencing), each of which can be modified independently of the others. When the user starts the programme these components are assembled and presented dynamically: the runtime engine looks for the XML parameter file and, finding it, loads the instructions and content specified in the XML. Some of these specifications are URL pointers to external resources such as images, animations, data traces, text and questions that are dynamically loaded in to the runtime programme as well.

Any one of the components can be changed independently of any of the others – images or text can be added or edited, sections of the CAL programme can be extended, the programme can be translated into any number of different languages or a different runtime engine could be used. In this situation, changes to the XML file would point at a different set of text and media assets as local circumstances dictate. While Flash currently provides the runtime engine of choice, the emergence of a new runtime technology in a few years would not create problems, as the XML sequence and the media assets are managed separately and would simply be plugged in to the new runtime engine. RECAL is also developing a web runtime engine.

The RECAL process

The RECAL process comprises a number of stages.

Disaggregation

RECAL is working with an existing set of CAL programmes developed by the author (DD) and now marketed through Sheffield BioScience Programmes (www.sheffbp.co.uk). These CAL programmes are pedagogically sound and it is important that the assets and learning design encapsulated in them are not lost simply because the runtime engine has become or is likely to become obsolete. Part of the RECAL process is to unlock the content and release the media assets such as images, diagrams and animations, information assets such as text, data and algorithms, and learning activity designs such as the structure, organisation and rules the CAL programmes contain. In some cases the CAL programmes are available in an uncompiled form, which allows direct access to these components; in others the components have to be extracted by taking screen-grabs, transcribing text and by recreating those components where no direct extraction is possible. Animated sequences, diagrams and photographs usually need to be recreated, either because the animation is inextricably tied to the particular authoring tool used or because they do not meet contemporary standards of quality. At the end of the disaggregation process we are left with: text (XHTML), images (JPEG), dynamic animations and images (SWF), questions (IMS QTI XML) and parameter files describing the sequencing (IMS SS XML).

Asset (learning object) management

Disaggregation of each CAL programme can release in excess of 100 assets, each of which needs to be managed independently of the others. Each asset is catalogued; this associated metadata describes its basic properties and includes file type, location, keywords, intellectual property rights, provenance, location, and size. The asset and its metadata are managed in a learning object repository, designed to specifically hold and manage learning object metadata (using the IEEE 1484 Learning Object Metadata standard).

Pedagogical design

The pedagogical design of the original CAL programmes is rendered in XML using IMS Simple Sequencing specification. IMS Simple Sequencing, in that it consists of ways to describe a learner's progress and interactions, is most like traditional CAL programmes, with sequencing rules covering aspects such as branching or conditional advancement.

Run-time engines

The run-time engine (an application that renders the presentation elements and manages user input) has to be able to read the XML parameter files, call down the appropriate resources from

the repository and provide the interface for the user. RECAL has decided to use Macromedia Flash for its standalone run-time shell but is also testing web rendering in plain text to support users with visual disabilities.

Authoring tools

The creation of easy-to-use, template-based authoring tools will enable non-technical teachers to assemble their own learning activities from the disaggregated assets. The goal is to provide teachers with an editor that lets them customise all aspects of the CAL programme for their local needs. This customisation may include the reorganisation, addition or removal of sections, images, text, or questions, changing things such as drug dose, drug names or electrical stimulation parameters. The tools could also be used to create different language versions. At its core, the authoring tool will create XML text files for the run-time engine to use and will use a simple graphic interface to support the majority of authors who will be unfamiliar with XML.

Results

RECAL pilot

A proof-of-concept pilot of this staged process was completed in the autumn of 2003. A Flash run-time was constructed that could dynamically load and render an external XML parameter file and the resources it identified. The pilot only involved one CAL programme, an English-language version of a programme originally developed for DOS, later rewritten for early versions of Windows using Assymetrix Toolbook version 2.0 and redeveloped again in 1998 for later versions of Windows using Macromedia Director. The programme was disaggregated and the XML was assembled in a text editor.

The pilot's short development cycle has raised a number of technical implementation issues:

templates for different types of screen/page types, which support different types of interaction, need to be developed, e.g. data display screens, multiple choice question screens, information screens.

Navigation often needs to be simplified and now uses common web-browser icons e.g. the Home button and the back/forward buttons closely resemble those used in Microsoft's Internet Explorer.

The tools are being developed and evaluated in-house with the purpose of tailoring the pilot application for non-technical teachers.

Supporting simulations of experiments

Multimedia CAL alternatives to using animals in teaching pharmacology and physiology are broadly of two types:

CAL simulations of pharmacology experiments typically present data traces for a number of key experiments designed by experienced teachers to meet identified learning goals. Students have limited control over the design of an experiment. These programmes often include significant background information, self-assessment questions designed to test accuracy of data recording, data interpretation, knowledge of underlying principles, etc. This approach would be analogous to the learner using a tutor-designed schedule to perform a series of defined experiments on a pharmacological preparation where the teacher defines the experimental parameters such as which drugs/drug combinations, drug-dose range to try to ensure that the experiment is successful and the learning goals are achieved.

CAL simulations of preparations that are algorithmically-driven and try to put the learner into a situation where he or she is provided with a pharmacological preparation and a number of test drugs. The learner must design experiments to meet certain learning objectives and must choose experimental parameters such as drug dose, drug combination and other experimental parameters.

Both approaches have merit and different learning objectives will be achieved. The majority of the programmes available to the RECAL project are computer simulations of experiments and much of the development has focussed on this approach. However, if RECAL is to have wider applicability, it needs to be able to handle algorithms that create simulated data. Although this has not as yet been tested, it should be possible for RECAL to handle more complex interactions such as algorithmic simulations.

Discussion

This paper has described the RECAL approach to extending the life of a number of existing computer-based alternatives. It has the potential to provide teachers and courseware developers with a mechanism for salvaging good pedagogical content from CAL programmes that have reached or are nearing technological obsolescence. The disaggregation process releases the components or assets and the educational designs that constitute these CAL programmes and makes them available for reuse via a repository into which all of the component learning assets are catalogued.

Abstracting the learning assets and the educational design from the run-time engine approach is also recommended as a method for new developments. The 3 year project is very much work in progress and is now at mid-point. The results of the proof-of-concept pilot are very encouraging and about 60% of the computer programmes available to the project have been disaggregated, yielding in excess of 1000 learning assets ranging in complexity from text, images and data traces to interactive diagrams and questions and animated graphics.

The development of easy-to-use, template-driven authoring tools will empower teachers to create their own learning activities rather then being reliant on the expertise of a courseware developer. These same tools will enable teachers to edit existing computer simulations, which have been developed using the RECAL methodology, to meet their local learning needs and circumstances.

Although the process has been developed specifically for an existing range of computer simulations, it could be applied to any multimedia CAL programme in any discipline. The opportunity that this approach offers to an institution to enable it to better manage valuable teaching and learning resources more effectively should also not be overlooked.

References

Dewhurst, D. G., Hardcastle, J., Hardcastle, P. T. and Stuart, E. (1994). Comparison of a computer simulation programme with a traditional laboratory practical class for teaching the principles of intestinal absorption. *Amer. J. Physiol. 267, (Advances in Physiology Education) 12 (1)*, S95-S103.

Fawver, A. L., Branch, C. E., Trentham, L. et al. (1995). A comparison of interactive videodisc instruction with live animal laboratories. *American Journal of Physiology: Advances in Physiology Education 259*, S11-S14.

Gruber, F. P. and Dewhurst, D. G. (2004). Alternatives to Animal Experimentation in Biomedical Education. *ALTEX 21 (Suppl. 1)*, 33-48.

Guy, J. F. and Frisby, A. J. (1992). Using interactive videodiscs to teach gross anatomy to undergraduates at The Ohio State University. *Academic Medicine 67 (2)*, 132-133.

Hughes, I. E. (2001). Laboratory Practicals in Pharmacology Teaching – do they meet the learning needs? *Trends in Pharmacological Sciences 22 (2)*, 71-74.

Kinzie, M. B., Strauss, R. and Foss, J. (1993). The effects of an interactive simulation on the performance and achievement of high school biology students. *Journal of Research in Science Teaching 30 (8)*, 989-1000.

Leathard, H. L. and Dewhurst, D. G. (1995). Comparison of the cost-effectiveness of a computer assisted learning programme with a tutored demonstration to teach intestinal motility to medical students. *Association for Learning Technology Journal 3 (1)*, 118-125.

Acknowledgements

Financial support for the RECAL project is kindly provided by The Lord Dowding Fund (NAVS), UK.

Correspondence to

Professor David Dewhurst
Assistant Principal (e-learning & e-health)
Director of Learning Technology
College of Medicine & Veterinary Medicine
University of Edinburgh
15 George Square
Edinburgh EH8 9XD, UK
e-mail: david.dewhurst@ed.ac.uk

The InterNICHE Policy on the Use of Animals and Alternatives in Education

Nick Jukes[1] *and Siri Martinsen*[2]
[1]InterNICHE, Leicester, England; [2]InterNICHE Norway, Oslo, Norway

Summary
The InterNICHE "Policy on the Use of Animals and Alternatives in Education" is a comprehensive document in 10 sections that addresses all aspects of work with animals and alternatives in life science education. The Policy presents guidelines to ensure effective and fully ethical acquisition of knowledge and skills. It includes a definition of alternatives in education and of harm, and presents individual policies on dissection, the sourcing of animal cadavers and tissue, work with live animals for clinical skills and surgery training, and ethical field studies. It also addresses the use of animals for the production of alternatives. While the ideal "replacement alternative" is defined as "non-animal" within the 3Rs philosophy of Russell and Burch (1959), the Policy highlights a shortcoming of this approach for education. Not only is there a requirement for some students to work with animals, animal tissue and clinical procedures in their education, there is widespread evidence of the ability to fully meet all teaching objectives in ways that are neutral or beneficial to individual animals and that do not involve animal experimentation or killing. As well as non-animal learning tools, like multimedia computer simulation, digital video, training models and mannequins, replacement alternatives also include the use of ethically-sourced animal cadavers for dissection and skills training and apprenticeship into clinical practice with animal patients. Definitions of "ethically-sourced" and of ethical educational opportunities within clinical work are included in the Policy, which demonstrates the possibilities for full replacement of harmful animal use in education.

Keywords: alternatives, animal experimentation, clinical, dissection, education, ethically-sourced, field studies, InterNICHE, teaching objectives, harm, policy

Introduction

The InterNICHE Policy on the Use of Animals and Alternatives in Education (Jukes and Chiuia, 2003) presents comprehensive guidelines in 10 sections that address all aspects of work with animals and alternatives in life science education. The Policy reflects the InterNICHE commitment to full replacement of harmful animal use whilst supporting effective acquisition of knowledge and skills.

Beginning with the InterNICHE Position Statement, a definition of alternatives in education and a definition of harm, the document presents individual policies on a range of issues including dissection, the sourcing of animal cadavers and tissue, work with live animals for clinical skills and surgery training, and field studies. It also addresses the use of animals for the production of alternatives.

Why is a policy necessary?

A policy provides a considered opinion on an issue or practice and through its guidelines can facilitate transformation towards best practice. It is important for organisations to have a clear position with respect to all aspects of their field of focus in order to effectively communicate their vision and practical strategy and to optimise the positive impact of their work. The InterNICHE Policy provides a detailed exploration of the ethical and practical issues inherent in the use of animals and alternatives, and envisages and facilitates the creation of a fully humane education so that the teaching objectives of practical classes can be met in the most ethical and effective ways.

The process of curricular design and the selection of learning tools in life science education involve choices by decision-makers that always have an ethical dimension. The Policy clarifies the ethical impact of the different uses of animals and alternatives, states what is and is not acceptable according to InterNICHE, and makes clear the possibilities for enhancing the quality of education.

The Policy provides a base for developing and modernising regulations and legislation concerning the use of animals and alternatives in education, from the level of a university department to national and international law. It also helps individual students and others to clarify and reflect on their ethical positions and encourages capacity-building in ethical decision-making and practice.

Contents of the Policy

The 10 sections of the Policy are:
1. Position statement
2. Definition of alternatives in education
3. Definition of harm
4. Policy on animal dissection
5. Policy on ethical sourcing of animal cadavers and tissue
6. Policy on other sources of animal cadavers and tissue

7. Policy on live animal use for clinical skills and surgery training
8. Policy on live animal field studies
9. Policy on the ethical use of live animals, animal cadavers and tissue for making alternatives
10. Policy on other use of live animals, animal cadavers and tissue for making alternatives

Each of the sections is briefly described below.

1. Position statement

InterNICHE supports a high quality humane education within the life sciences and the use of alternatives to meet teaching objectives. InterNICHE is against all harmful use of animals in education, including the harming and killing of animals for their cadavers and tissue, for live experimentation and skills training, for ethology and field studies and for making alternatives.

2. Definition of alternatives in education

Alternatives are humane educational aids and teaching approaches that can replace harmful animal use or complement existing humane education. Alternatives may be non-animal alternatives or approaches that involve neutral or beneficial work with animals. They comprise:

- Film and video
- Models, mannequins and simulators
- Multimedia computer software and virtual reality (VR)
- Ethically-sourced animal cadavers and tissue
- Clinical work with animal patients
- Student self-experimentation
- *In vitro* labs
- Field studies

Many of the above will be recognised as part of educational best practice, and in some countries, humane approaches within certain disciplines are tradition. "Alternative" teaching approaches – by tradition or by modern choice – are therefore often the norm. For example, the training of veterinary students involves direct experience with real patients in all primary veterinary colleges, and it is undertaken as part of the provision of clinical care by faculty clinicians or private practitioners. However, in some colleges and some countries, students still typically use laboratory animals for clinical skills and surgery training and may have little or no access to the learning opportunities available with clinical work on patients. Developing significantly greater student access to such clinical work is therefore an "alternative" to the use of laboratory animals. Similarly, other tools and approaches listed above are used widely by teachers but can all be considered as alternatives where harmful use of animals continues.

The majority of life science students will enter professions that do not involve work with animals, and well-designed combinations of non-animal learning tools can successfully meet the teaching objectives in their practical classes. This leaves the privilege of working with animals to students of veterinary medicine, zoology and other fields in which there is a requirement for students to be familiar with animals, animal tissue and clinical procedures. Such experience can be achieved in ways that are neutral or beneficial to the well-being of individual animals and that do not involve laboratory animal experimentation or killing.

As the ideal "replacement alternative" is defined as "non-anima"' within the 3Rs philosophy of Russell and Burch (1959), the Policy therefore also highlights a shortcoming of this approach for education: despite the widespread success and further potential for replacement of harmful animal use by non-animal alternatives, the latter are not sufficient for full knowledge and skills acquisition in certain areas of education. Veterinary surgery is one example where thorough training with non-animal alternatives must be followed by hands-on experience with living animals. However, this does not suggest a necessity for conventional training using laboratory animals. Instead, it suggests that approaches that are neutral or beneficial to individual animals, such as clinical work with patients, are the ideal replacement alternatives. This full replacement of all *harmful* animal use, rather than of all work with animals, is the appropriate solution to many of the pedagogical, ethical and practical challenges facing life science education. The Policy has a significant focus on the ethical use of animals in order to address this issue.

3. Definition of harm

Integral to discussions about harmful animal use and alternatives is the definition of harm. While there may often be agreement on whether a certain practice or procedure is harmful to an animal and efforts are made to assess the degree of that harm, disagreements and individual and cultural differences in such assessments remain. For example, is killing a form of harm? Is an experiment performed under general anaesthetic harmful? Is isolating an animal from others of his/her species harmful?

According to the Policy, harm comprises any action, deliberate or otherwise, that impinges on an animal's current and future well-being by denying or limiting any of the following freedoms:

- Freedom to live
- Freedom to express full natural behaviour
- Freedom to be part of a social structure and ecosystem
- Freedom from hunger and thirst
- Freedom from discomfort
- Freedom from pain, injury and disease
- Freedom from fear and distress

This is a strict definition of harm, but it reflects its serious nature. Moreover, case studies of full replacement of harmful animal use demonstrate that the exacting demands of the Policy can indeed be met (Kumar, 2003; Smeak, 2003).

Harm caused to an animal within education is only acceptable when it is the unavoidable consequence of action taken to benefit the individual animal, and in certain circumstances when the action is taken to benefit the species or to produce an alternative, given that the harm inflicted is only brief and minor. In these cases, cost-benefit analyses concerning harm and potential benefit to the individual animal, the species and to other animals should be conducted.

4. Policy on animal dissection

Animal dissection can be a useful tool for knowledge and skills acquisition and may encourage an appreciation of life, when all

8 conditions described in this section of the Policy are met. Three of these conditions are described below:

4.1 The animal cadaver is ethically-sourced (see section 5 below).

4.2 The dissection is performed at university level and no lower.

4.3 The dissection is relevant for the student's career.

The majority of dissections performed at universities are of animals killed for that purpose, and such use is not acceptable.

5. Policy on ethical sourcing of animal cadavers and tissue

InterNICHE is against all harmful use of animals in education, including the harming and killing of animals for sourcing of cadavers and tissue. As illustrated by the dissection examples above, however, such material may provide a useful resource for knowledge and skills acquisition. The use of an animal cadaver or tissue is acceptable when it is ethically-sourced, and it may be recognised as such when all 9 conditions in this section are met. The first three state:

5.1 The animal was not captured, bought, bred, kept, harmed or killed to provide the cadaver or tissue.

5.2 The animal was wild, stray, or a companion animal before death.

5.3 The animal died from natural causes or an accident, or was humanely euthanised secondary to natural terminal disease or serious non-recoverable injury.

Animal cadavers and tissue are usually obtained from sources where animals suffer harm or are killed, such as animal breeders, research facilities, some animal shelters, farms, slaughterhouses and sporting events. The Policy does not consider material from these sources, including so-called "waste" or "surplus" material, to be ethically-sourced: its ethical nature has been compromised or negated by the harming, killing and/or marketing of the animal at some stage of his/her life.

The condition in this section and elsewhere in the Policy that the animal must be a wild, stray, or companion animal is to ensure that the animal was free-living in origin rather than bred and/or kept for use. The reference to euthanasia refers to the true meaning of the word: ending the life of an animal that is suffering from terminal illness or an incurable condition, based on the interests of the animal. Despite its necessity, euthanasia is referred to in the Policy as harmful, because taking of a life is a significant form of harm.

The following examples of source and use of animal cadavers and tissue further illustrate the Policy:

First, anatomical dissection of frogs collected from the wild and sold by a biological supply company: As the frogs were captured and killed, the cadavers are not ethically-sourced and such use is not acceptable.

Second, surgical dissection by veterinary students of companion animal cat and dog cadavers sourced from a veterinary teaching hospital: If the animals died from natural causes or in accidents or were euthanised for the reasons stated in 5.3 (above), then the cadavers are ethically-sourced and such use is acceptable. Dogs killed in shelters due to inability to rehome them are not ethically-sourced because they do not meet the conditions of 5.3.

Third, mitochondria from the livers of rats sourced from research laboratories: As the animals are laboratory animals and not wild, stray or companion animals, and moreover were killed rather than euthanised, the tissue is not ethically-sourced and such use is not acceptable. Tissue from companion animal rats is ethically-sourced if all conditions are met.

6. Policy on other sources of animal cadavers and tissue

In non-ideal circumstances, cadavers and tissue from species of animal that are less common as wild, stray or companion animals may be hard to source ethically. This may be the case in some countries at certain times. Deriving animal cadavers and tissue from other sources such as research facilities, slaughterhouses or farms is therefore an acceptable compromise when all 8 conditions in this section are met. The first three state:

6.1 Animal cadavers or tissue are genuinely required for practical work or for making an alternative, and no ethically-sourced and appropriate material is available.

6.2 The animal was not captured, bought, bred, kept, harmed or killed to provide the cadaver or tissue.

6.3 The animal has died from natural causes or an accident or was humanely euthanised secondary to natural terminal disease or serious non-recoverable injury; or the cadaver or tissue is destined for disposal or has been abandoned by the animal.

Genuine waste from the above sources is therefore acceptable under certain conditions. Material other than genuine waste is not acceptable, because it helps to create a demand for animals from such sources, which contributes to further harming or killing. So-called "surplus" animals deriving from research facilities and sometimes used in education are an example of what is not acceptable.

7. Policy on live animal use for clinical skills and surgery training

The use of live animals in the clinical setting is an integral part of knowledge and skills acquisition for veterinary students. It is acceptable when all 14 conditions in this section are met. First, in contrast to the conventional instrumental use of animals, the Policy states:

7.1 Opportunities for clinical skills and surgery training are built around the needs and well-being of individual wild, stray and companion animal patients, and healthy companion animal volunteers.

Such an approach is described in detail by Rasmussen (2003).

Although much clinical skills and surgery training involves no harm to animals, further conditions in this section explore the issue of harm and the performance of terminal procedures.

7.5 Harm caused to an animal patient during a clinical procedure and/or treatment is acceptable when it is the minimum harm necessary for successful work aimed at healing the animal; and in certain circumstances during procedures involving an animal that is suffering from natural terminal disease or serious non-recoverable injury; or when it comprises the act of humane euthanasia.

7.6 Clinical skills and surgery training that involves a ter

minal procedure is acceptable only when an animal is suffering from natural terminal disease or serious non-recoverable injury; and for whom a decision to euthanise has been made by a qualified veterinarian with the consent of the animal's guardian (if any), based on the interests of the animal and not motivated by practical or financial interests.

7.7 Harm caused during an invasive and/or terminal procedure on an animal that is suffering from natural terminal disease or serious non-recoverable injury is acceptable only when the harm is not subjectively experienced by the animal; and when it comprises the act of humane euthanasia.

Other conditions in this section help ensure that live animal use is acceptable: appropriate training and competence of the instructors; supervision and sufficient skills mastery of the student; respect for the student's ethical position; respect for the animal; high welfare, health and safety standards; and open ethical discussion in the classroom. All are necessary conditions for work with live animals in the clinic. Some of these conditions are repeated in other sections of the Policy.

8. Policy on live animal field studies

The educational study of free-living wild or stray animals is a valuable experience, acceptable when all 14 conditions in this section are met. The first four state:

8.1 Opportunities for field studies are built around the needs and well-being of individual wild and stray animals, animal species, and the ecosystem.

8.2 The animal is not captured, bought, bred, kept, harmed or killed for the purpose of the study, except for capture and/or harm in certain circumstances that are beneficial to the individual animal, species or ecosystem.

8.3 Field studies should cause zero or minimal disturbance to an animal, his/her social structure and the ecosystem; or have a beneficial impact on an animal, species or ecosystem.

8.4 Capture and/or harm caused to an animal are acceptable only when the animal is a patient, or will benefit from a clinical procedure; and in certain circumstances for the benefit of the species or ecosystem.

The potential for ethical field studies is described by Bekoff (2005).

Other conditions in this section clarify that harm caused to an individual animal for the benefit of the species or ecosystem is acceptable only if minor and temporary; and state that field studies should avoid threatened species and ecosystems unless the likely benefits outweigh the costs; and refer back to sections of the Policy on sourcing of animal cadavers and tissue, on live animal use for clinical skills and surgery training, and on the use of live animals, animal cadavers and tissue for making alternatives (see below), when field studies involve these aspects.

Two examples illustrate the Policy. First, temporary capture and tagging of a wild animal for ethology studies that will support the protection of the species and the ecosystem: As long as the individual animal's future well-being is not jeopardised by the capture and tagging, this is acceptable. Second, temporary capture of feral cats to study lineage and genetics: This is only acceptable if the individual animals will also benefit from clinical attention.

9. Policy on the ethical use of live animals, animal cadavers and tissue for making alternatives

Animals are also used for the production of alternative tools, such as videos or software for virtual dissection and virtual experimentation. Typically, this involves harming or killing animals, and such use is not acceptable. Data generation in virtual experimentation can in some cases be achieved using mathematical algorithms. However, in other cases the use of a live animal or an animal cadaver or tissue for making an alternative may be considered necessary. Such use is acceptable when all 7 conditions in this section are met. The first two state:

9.1 An alternative for the practical does not already exist or is not practicably available.

9.2 The animal is not captured, bought, bred, kept, harmed or killed for the purpose of making the alternative, except for harm and/or euthanasia in certain circumstances during procedures involving invasive and/or terminal live animal use.

Other conditions re-iterate sections of the Policy that address the use of animals, and apply them to the context of making alternatives. These include sections of the Policy on sourcing of animal cadavers and tissue, on live animal use for clinical skills and surgery training, and on live animal field studies, when these aspects are used to make alternatives.

One example illustrates the Policy, particularly the issue of harm in 9.2. Filming a terminal surgery on an animal that is facing euthanasia in order to make a needed alternative: If the decision to euthanise was made by a qualified veterinarian, based on the interests of the animal, and no additional harm is caused by the surgery, then such use is acceptable. It would not be acceptable to perform the same procedure on a terminally ill dog if it causes him/her additional harm, nor on a healthy shelter dog.

10. Policy on other use of live animals, animal cadavers and tissue for making alternatives

Under non-ideal circumstances the ethical production of alternatives may occasionally prove difficult. This section describes how the non-ideal use of and non-ideal sources of live animals, animal cadavers and tissue for making alternatives may under certain conditions provide an appropriate solution to this ethical challenge.

Deriving live animals for use in invasive and/or terminal procedures and animal cadavers and tissue from other sources, such as research facilities, slaughterhouses or farms, is therefore an acceptable compromise for the purpose of making an alternative when all 13 conditions in this section are met. The first three state:

10.1 An alternative for the practical does not already exist or is not practicably available.

10.2 The animal is genuinely required for making the alternative and no ethical source of a live animal or animal cadaver or tissue is available.

10.3 The alternative to be made will replace harmful animal use in education, and will be widely available for students to use.

Other conditions in this section state:

10.5 If live animal use is required, priority is given to the sourcing and use of an animal when he/she is suffering from natural terminal disease or serious non-recoverable injury and for whom a decision to euthanise has already been made by a qualified veterinarian with consent of the animal's guardian (if any), based on the interests of the animal and not motivated by practical or financial interests.

10.6 If an animal could be recovered and rehomed, then he/she should be recovered and rehomed and not used for a terminal procedure or one that will necessitate euthanasia.

10.7 All sourcing of a live animal and invasive non-terminal live animal use should result in some direct or indirect benefit for the animal, such as being saved from euthanasia, being neutered during a procedure, and being recovered and rehomed.

10.8 All invasive live animal use should cause zero additional harm to the animal and should not jeopardise the animal's future well-being except in certain circumstances during procedures involving an animal that is suffering from natural terminal disease or serious non-recoverable injury.

10.9 Harm caused during an invasive and/or terminal procedure on an animal is acceptable only when the harm is not subjectively experienced by the animal; and when it comprises the act of humane euthanasia.

Two examples illustrate the Policy. In both examples, it is assumed that no suitable alternatives already exist, that the planned alternative will bring about replacement, and that no ethical source of a live animal is available.

First, filming a surgical procedure performed on a terminally ill pig that is facing euthanasia at a farm: If the decision to euthanise was made by a qualified veterinarian, based on the interests of the animal and no additional harm is caused by the surgery, then such use is acceptable.

Second, filming a surgical procedure on a healthy pig at a farm: Although the procedure is not necessary for the individual pig, if no additional harm is caused by the surgery, then such use is acceptable if the pig is then saved from slaughter and rehomed. Such use is not acceptable if the pig is given to slaughter afterwards.

Summary

The InterNICHE Policy on the Use of Animals and Alternatives in Education presents comprehensive guidelines to support ethical and effective acquisition of knowledge and skills. Its strict nature reflects a strong commitment to the quality of life science education and makes clear the opportunities associated with alternatives and the freedom available to teachers to apply modern technology and progressive ethical thought to the learning process. The Policy helps shift the focus of the discussion from animal experimentation to that of best practice education and the tools and approaches to implement best practice. Integrating the Policy into regulations and legislation will further support modernisation of life science education.

Viewing the Policy

The Policy was first published in the book *From Guinea Pig to Computer Mouse: Alternative Methods for a Progressive, Humane Education* (Jukes and Chiuia, 2003). The Policy is an evolving document and the latest version is published on the InterNICHE website www.interniche.org. Comments from readers are welcome.

References

Bekoff, M. (2005). Field Studies and Animal Models: Towards Non-invasive Approaches in Zoology Research and Teaching. In N. Jukes and S. Martinsen (eds.), *Alternatives in the Mainstream: Innovations in life science education and training*. Proceedings of the 2nd InterNICHE Conference, 2005 May 12-15; Oslo, Norway. Leicester, UK: InterNICHE; forthcoming.

Jukes, N. and Chiuia M. (2003). *From Guinea Pig to Computer Mouse: Alternative Methods for a Progressive, Humane Education*, 2nd ed. Leicester, UK: InterNICHE.

Kumar, A. (2003). Client donation program to meet the needs of veterinary medical education: Alternatives to healthy animal sacrifice. In N. Jukes and M. Chiuia (eds), *From Guinea Pig to Computer Mouse: Alternative Methods for a Progressive, Humane Education*, 2nd ed. (107-116). Leicester, UK: InterNICHE.

Rasmussen, L. M. (2003). A pedagogically sound innovative and humane plan for veterinary medical education. In N. Jukes and M. Chiuia (eds.), *From Guinea Pig to Computer Mouse: Alternative Methods for a Progressive, Humane Education*, 2nd ed. (125-133). Leicester, UK: InterNICHE.

Russell, W. M. S. and Burch, R. L. (1959) *The Principles of Humane Experimental Technique*. Wheathampstead, UK: Universities Federation for Animal Welfare. Reprinted by UFAW, 1992.

Smeak, D. D. (2003). Ethical surgery training for veterinary students. In N. Jukes and M. Chiuia (eds), *From Guinea Pig to Computer Mouse: Alternative Methods for a Progressive, Humane Education*, 2nd ed. (117-124). Leicester, UK: InterNICHE.

Correspondence to

Nick Jukes
InterNICHE
42 South Knighton Road
Leicester LE2 3LP
England
e-mail: coordinator@interniche.org
www.interniche.org

Siri Martinsen
c/o NOAH
Osterhausgt. 12
0183 Oslo
Norway
e-mail: siri.martinsen@bredband.net

Internationalising Alternatives in Higher Education

Nick Jukes
InterNICHE, Leicester, England

Summary
InterNICHE has been working internationally to promote and implement alternatives in higher education for 17 years, facilitating the replacement of harmful animal use and building a broad network with contacts in over 50 countries. From the InterNICHE experience, successful international work requires qualities and practices from organisations that include: a bold and positive vision, a specific focus and an awareness of the links between issues; a commitment to pro-actively catalyse sustainable change and create win-win solutions; the design of organisational structures conducive to participatory democracy, alliance building and the organic growth of the network; the practice of solidarity and support for local initiatives rather than empire building; and the provision of resources and training for action and capacity building. The presentation will draw on examples of InterNICHE projects such as the production and multi-language translations of printed, video, and website resources; the Alternatives Loan System for trial of software, mannekins and simulators anywhere in the world; the international Humane Education Award for local development and implementation of alternatives, including freeware; support for student conscientious objectors; and conferences, outreach visits, and training in alternatives for teachers. The challenges met within such work will also be explored, and suggestions of how to overcome them will be given.

Keywords: alternatives, animals, education, freeware, grant, internet, InterNICHE, library, students, training, translation

Introduction

The International Network for Humane Education (InterNICHE) was founded in 1988 by student conscientious objectors, animal welfare scientists and anti-vivisectionists from Western Europe. It has since expanded to become a global network with contacts in over 50 countries and a diverse range of resources, projects and achievements, all focusing on the implementation of best practice teaching approaches using alternative methods and the practical replacement of harmful animal use. The growth of this unique network has been achieved through much voluntary commitment and with minimal financial resources. This paper explores the qualities and practices possible from organisations that can facilitate successful international work, using examples of InterNICHE philosophy, activity and organisational structure. It also addresses the challenges met within such work, and gives suggestions of how to overcome them.

Philosophy and vision

InterNICHE pro-actively catalyses progressive and sustainable change in life science education. The philosophy illustrated below informs this campaigning action. Philosophical consideration helps explore the ethics of an issue in depth, and ensures relevance and consistency of argument in an evolving field. It also informs organisational structure, decision-making and collaborative action, as well as the choice and design of projects and resources.

The InterNICHE vision is one of a fully humane education, brought about through full replacement of harmful animal use and the implementation of progressive learning tools and approaches. This bold and positive vision reflects the commitment and confidence of the international network of volunteers and collaborators. It is supported by the growing evidence of replacement that has already been achieved at universities across the world. The InterNICHE *Policy on the Use of Animals and Alternatives in Education* (Jukes and Chiuia, 2003) details in 10 sections the practical expression of this philosophy and vision. It provides comprehensive guidelines to inform curricular design, regulations and legislation.

Clarity of vision requires knowing the perspectives from the diversity of countries and world cultures. Awareness of difference as well as the shared value base and commonality of experience has supported the vibrant organic growth of the international network and of the broader humane education movement. A commitment to consensus decision-making and participatory democracy has further supported this.

InterNICHE campaigners keep a specific focus on alternatives in education whilst staying aware of the broader impact of curricular transformation in the scientific, ethical, pedagogical, social, and economic spheres. Working with this holistic awareness creates more opportunities for alliances to be found, avoids shifting problems from one sphere to another with an illusion of success, and allows for genuine progressive change with multiple positive impacts to be achieved. InterNICHE believes that win-win solutions to the ethical conflict surrounding harmful animal use are almost always possible, with replacement benefiting students, teachers, animals and the life sciences in general.

InterNICHE also has a strong commitment to critical thinking and ethical literacy, both of which are crucially important to education and true scientific endeavour. Critical thinking is the abil-

ity to question assumptions and ask for evidence, for example of the claimed but elusive advantages of harmful animal use over alternatives. Such a commitment also includes finding creative solutions to genuine challenges – such as exploring the practicalities of securing ethically-sourced animal cadavers as an alternative, so that even hands-on dissection can be achieved in a fully ethical way.

Creating sustainable change

InterNICHE aims to create sustainable change through presenting the vision of full replacement and focusing on practical ways in which the vision can be implemented on the ground, taking account of each situation's unique opportunities and challenges. This is done through the win-win approach, solidarity and support for local initiatives, and provision of resources and other practical support.

InterNICHE campaigners work with teachers, students and other stakeholders. In some cases consensus can be reached easily and change implemented rapidly. InterNICHE provides one of the few forums where both abolitionists and animal experimenters can meet to discuss and search for common ground upon which to build some progressive change involving replacement. The existence of such a medium for communication is rooted in respect for people, inclusivity and empowerment: a recognition that everyone can be an agent for progressive change.

When such a co-operative process has been precluded by teachers, however, campaigning may involve actively challenging harmful animal use and the denial of student choice, peacefully and responsibly. Even during this, however, the overall aim is still on the positive: addressing best practice teaching through the most effective and ethical ways of gaining knowledge and skills.

The provision of resources to create more opportunities for achieving win-win solutions and to support people to make changes themselves are crucial aspects of achieving sustainable change. The InterNICHE Alternatives Loan System, offering alternatives for free loan to all countries, has given teachers opportunities to try out and become familiar with the diversity and quality of a range of existing products. "Hands-on" experience of alternatives and opportunities to trial products in advance of purchase are often essential steps leading to effective implementation. Other support includes the provision of training within countries, with alternatives that are relevant to the local curricula; a Humane Education Award, co-judged by National Contacts, that has supported the purchase, development and implementation of alternatives; and multi-language translations of information material, including the InterNICHE book, video and website described below. Many other smaller local initiatives are also supported.

Working with National Contacts and collaborators to support their own initiatives leads to change coming from within rather than being imposed from outside, and the change is therefore more likely to be fully relevant as well as sustainable. It is also likely to be more effective, as it is rooted in and responsive to local situations and values, and draws on existing local skills, knowledge and experience rather than ignoring them. With this approach of solidarity and partnership rather than empire building, InterNICHE can facilitate decentralised activity across the world.

Provision of resources and practical support

Some of the resources and practical support made available to facilitate sustainable change are described below.

Information

Information resources are necessary for teachers and others to make informed choices regarding curricular change and the right combinations of alternatives for the location. Gathering information on animal use and the ongoing replacement at universities is one task, and National Contacts may be active doing this. Provision of information is another task, and the network researches, provides and distributes such information internationally.

The InterNICHE book *from Guinea Pig to Computer Mouse* (Jukes and Chiuia, 2003) provides background information including studies of curricular design and assessment of alternatives, case studies written by university heads of department who have fully replaced harmful animal use, and links to many further resources. Full details of over 500 alternatives, including description, specification and source, comprise the bulk of the book, which is available in an increasing number of languages. The complete text of the book will is now available on-line, with plans for a searchable and evolving database. On-line databases on alternatives are increasingly being specified as "required visiting" before teachers and researchers apply to animal ethics/animal care and use committees, as part of the moral and legal burden of proof on teachers that they have investigated "all possible alternatives" to animal procedures.

The InterNICHE website at www.interniche.org provides news and information about advances in life science education, arguments for the implementation of alternatives, student testimonies and advice on conscientious objection, and details of the resources that InterNICHE offers. It provides links to other organisations and their resources, and lists the latest contents of the Loan System, with links to producers. Free to download is the 33-minute InterNICHE film *Alternatives in Education* (1999) which features interviews and demonstrations of alternatives by teachers, and the sound files from the 2nd InterNICHE Conference held in Oslo (2005). The site has an increasing number of visitors, and was approaching 1.5 million hits and nearly 100 GB of download for 2005. Translations of the site into new languages now include Portuguese, Polish, Russian and Arabic, amongst others, providing the first on-line alternatives texts in some languages. The site is currently being redesigned to comprise many new facilities, in particular a large degree of interactivity and accessibility, and more on-line resources. Regional variations of selected content will honour the cultural diversity relevant to the issue and within the network, providing an appropriate degree of localisation.

Support for conscientious objectors

Support for student conscientious objectors is crucially important in the face of some teachers' emotionally charged opposition to humane science, and the threat of academic or psychological penalty suffered by students. Information and practical advice on how to object, step-by-step, along with testimonies from student conscientious objectors, are provided at the InterNICHE website. The shared experiences of other students who have been through similar situations, and the community of objectors and campaigners that InterNICHE has helped build can give power to those objecting and sustain them through the process towards success.

Some National Contacts are or have been conscientious objectors at the forefront of pushing for curricular change; in some cases they are the first in their country to graduate using alternatives only. The role of conscientious objection in creating change is clearly illustrated by the example they have set within their discipline of what is possible, and by the alternative tools and approaches they may have helped implement for wider student access. The need for conscientious objection and the challenges facing such students when some teachers refuse to explore win-win solutions to ethical conflict illustrates the limits to students' freedom of conscience that are imposed within life science education by compulsory harmful animal use. The penalties sometimes imposed on critical thinking, ethically literate students who conscientiously object suggest that education is not as accessible as it should be. The life sciences need to decide whether a student's questioning of the orthodoxy, commitment to ethics and demand for educational best practice are more important than a teacher's attachment to convention. The latter have unlimited academic freedom to develop progressive humane ways of teaching.

Alternatives Loan System

InterNICHE built a large library of alternatives during 2001/2002 to support practically the process of replacement worldwide. This evolving Loan System is co-ordinated from Slovenia, with alternatives available for free loan to all countries worldwide. Borrowers request items and sign a Borrowers Agreement Form guaranteeing that they will not copy software and will return the items when requested and in good condition. They pay only the return shipping costs.

Over 100 CD-ROMs, videos, simulators and training mannekins are included for their pedagogical value and potential for replacement. The project has made over 200 loans to 40 countries, comprising over 4000 usages of individual alternatives, since its establishment. Borrowers include teachers, students, animal ethics committees, government ministries, organisations and campaigners. The loans have successfully given access to alternatives where none or little existed before, provided a resource for demonstrations at conferences, outreach tours and training, and supported the work of campaigners by providing a powerful international resource. As a tool for facilitating implementation, the value of the Loan System is indicated by a number of positive results: significant teacher use and the high number and wide geographical range of loans, positive feedback on the resource from borrowers, subsequent purchase and implementation of products, and direct replacement of harmful animal use.

Small-scale "micro-Loan Systems" have been established in Brazil, Russia, Ukraine, India and Japan, providing local resources under the management of National Contacts. These facilitate ease of borrowing by avoiding international shipping costs and by catering for particular cultural challenges and opportunities. These seed projects of the much larger international Loan System illustrate how much can be done with seed funding to support small-scale but highly effective and sustainable projects that are designed to facilitate replacement of harmful animal use. The establishment of new resources regionally and in different countries is to be encouraged.

Outreach and training

Outreach comprising individual presentations and seminars held at universities across the world can also be highly appreciated by teachers and students, particularly with visiting international speakers. Such widespread work has reached many new audiences. National Contacts may participate in or organise events or speaking tours across their country. Recent nationwide tours have taken place in Brazil, Czech Republic, Norway, Russia, India and Japan, and typically also involve one-to-one meetings with university faculty. Display and demonstration of alternatives are also made at these events.

To maximise the impact of such visits, InterNICHE has organised a number of alternatives training seminars over the years to impart skills that will support effective implementation. Such training comprises more than demonstrations of alternatives, and instead employs the expertise of local trainers who present in detail the content and potential of a range of alternatives, chosen according to the local curricula and situation. Alternatives from the Loan System are typically used, and training may involve collaboration with producers. In August-September 2004 over 400 university teachers were trained in alternatives and animal welfare at a series of seminars in 10 cities across India, in a project organised by InterNICHE in conjunction with the World Society for the Protection of Animals (WSPA) and many committed individuals and local organisations from across the country. The project involved teams of teacher trainers, and was the first of its kind worldwide that has provided training at a national level to such a large number of delegates.

The Multimedia Exhibition at this 5th World Congress was also organised by InterNICHE, with National Contacts and collaborators as trainers. Future training is planned for other conferences and for outreach tours in Latin America, North Africa and the Middle East.

Humane Education Award and freeware production

Life science courses across the world often show great similarity, and this is reflected in the practical classes and the animal use employed. Moreover, many individual alternatives are suitable for widespread implementation in such diverse locations. Despite this commonality and potential, however, specific local needs and other issues concerning alternatives, such as language and other cultural aspects, cost, tailorability, and a sense of ownership all play important roles. To address this, since 2002 InterNICHE has offered an annual Humane Education Award of 20,000 Euro to support multi-local development and implemen-

tation of alternatives. Supported by Proefdiervrij, this grant program is targeted at teachers or others who could bring about replacement through the production of new alternatives or the purchase and implementation of existing products. Submitted projects are judged on their potential to replace harmful animal use, potential pedagogic effectiveness, overall ethical design, commitment to open source, and other factors.

The Award has been focused regionally, beginning with South-Eastern Europe and then India, and is now global. Examples of projects funded include the following: First, the production of a freeware pharmacology compilation software in India, with over 3000 CDs distributed directly to teachers in the country. At just 3 universities, a total of 1600 animals were directly replaced by the software, so a much higher degree of replacement is likely to have been achieved nationally. Second, the production of a dual language physiology freeware in Romania and the establishment of a multimedia lab using reconditioned computers. The annual use of 1000 animals has been replaced at the department where the software was produced. Both freewares have also been distributed worldwide by National Contacts. A translation into Russian is also under production: free alternatives developed in Eastern Europe have great potential across the former Soviet Union due a strong similarity between courses in the region, and the financial difficulties facing universities. Third, the implementation of an advanced self-experimentation apparatus in Croatia, which as well as the replacement achieved has enabled new physiology practical classes to be established that were not possible with the animal labs. Fourth, the purchase of a range of Indian-made anatomy models for donation to an Indian university, with cow models distributed to many veterinary colleges across the country. This has contributed towards replacement of killed buffalo calves and of the painful embalming methods used.

The challenges

With some of the international co-ordination being done in the West – the root of much historical and contemporary empire building – it is important for the organisation to be mindful of potential cultural insensitivity in its practice. As well as the solidarity and partnership approach described earlier, the semi-autonomous National Contacts are responsible for much of the activity in their country and can ensure that the approach taken there is appropriate and effective. The Committee of National Contacts and the rotating Core Group are also spread right across the world, bringing diverse perspectives to planning and decision-making. The multicultural and internationalist experience of those involved also plays a role: some have travelled extensively and have a history of solidarity and partnership work. As part of the commitment to full inclusivity and equal opportunities, the challenges facing National Contacts and other campaigners from "developing" and other soft currency countries, particularly those relating to funding, access to visas for travel, and access to alternatives, always inform the decisions made and projects undertaken. This is also true for the broader international network that involves teachers, students and producers.

Occasionally campaigners may hold assumptions about or not appreciate the challenges faced by colleagues living in very different circumstances, but often this is simply a matter of exposure. The network itself creates opportunities to develop an internationalist awareness that appreciates common experience and purpose on the one hand, and diversity and difference on the other. The 2nd InterNICHE Conference had representatives from 32 countries, and internal meetings are equally multicultural. Sharing of experience and skills is an integral part of InterNICHE practice, both informally and with the specific purpose of empowerment, for example through training and the execution of projects. Similarly, the InterNICHE culture of consensus decision-making, wherever possible and appropriate, may be unfamiliar where an individual's personal preferences are more hierarchical or where they previously experienced little decision-making power.

There is always an element of transient involvement in the movement from students, which some may see as problematic. If InterNICHE has provided inspiration and resources to support students to question the orthodoxy, conscientiously object and strive to implement alternatives during their education, then the transience is of no matter. Indeed, some individuals may contribute after graduation, for example as professionals pushing for alternatives, informed by their own positive or negative experience in education. Difficulties caused by changes in active contacts are minimised by having available to the public a range of information and other resources, and by a commitment to empowerment and training, so that stability and continuity is maintained through open access and best use of resources and skills. Nevertheless, continuity can bring great benefits, and is essential in some organisational roles.

Structure and process evolve over time, and need constant review to ensure that they maximise the opportunities for participation and empowerment, and minimise those for bureaucracy. The Co-ordinator and Core Group perform much of the day-to-day international work, with autonomy for decision-making where appropriate; they are answerable to the Committee. Much general decision-making in InterNICHE is made by National Contacts at committee meetings and conferences, organised when finances allow. With a Committee based in nearly 40 countries, meetings can be a challenge both practically and financially. E-mail lists and chat are used at other times, and virtual meetings are being explored as a logistical solution to meeting in person. Networking opportunities at other events are used to the full, and regional meetings are also planned.

The rapid growth of the organisation, network and indeed movement for humane education has brought other challenges. More funds are needed, and though positive by being rooted in success, the challenge of work overload in the organisation can be difficult. The allocation of new roles within the organisation is one approach used to solve it. These include establishing new international roles, including regional co-ordinators, and encouraging National Contacts to build local groups. The role of named collaborators, including individuals or specific university departments, is being established to bring new people into the network and create non-executive roles with co-responsibility for specific projects. The establishment of the micro-Loan

Systems of alternatives is another aspect of ongoing decentralisation of resources that accommodates the growth and meets the needs of the network. And the restructuring of the InterNICHE website for more interactivity and automation allows for easier translation into new languages and for active participation by a broader network of individuals as the movement globalises further.

References

Alternatives in Education. Video (1999). Leicester, UK: InterNICHE. Also available at www.interniche.org/video.html.

Conference on-line. InterNICHE, 2005. www.interniche.org/2005conference/online.html.

Jukes, N. and Chiuia, M. (2003). *From Guinea Pig to Computer Mouse: Alternative Methods for a Progressive, Humane Education, 2nd ed.* Leicester, UK: InterNICHE. ISBN 1-904422-00-4.

Jukes, N. and Chiuia, M. (2003). Policy on the Use of Animals and Alternatives in Education. In N. Jukes and M. Chiuia, *From Guinea Pig to Computer Mouse: Alternative Methods for a Progressive, Humane Education, 2nd ed.* Leicester, UK: InterNICHE. Updated Version 2b available at www.interniche.org/policy.htm.

Correspondence to

Nick Jukes
InterNICHE
42 South Knighton Road
Leicester LE2 3LP
England
e-mail: coordinator@interniche.org
www.interniche.org

Workshop 1.3
Education in animal alternatives

Alternatives to the Use of Laboratory Animals in Veterinary Education

Vera Baumans
Department Animals, Science & Society, Division Laboratory Animal Science, Utrecht University, Utrecht, The Netherlands
Veterinary Resources, Karolinska Institutet, Stockholm, Sweden

Summary
About 75 million vertebrates are used worldwide per year for experimental purposes, of which 10 million within Europe. On average, about 2% is used for education and training purposes.
The basis of legislation on the use of animals for experimental purposes is the Three Rs principle of Russell and Burch: replacement, reduction and refinement.
Most legislation contains provisions to protect the animals such as the definition of legitimate purposes for animal use, competence of scientists and animal staff, the use of alternatives and prevention of unnecessary pain and distress. One of the legitimate purposes is the use of animals in education and training, only permitted when the objective cannot be achieved by comparable effective audiovisual or any other suitable methods (Art 25, ETS123).
In many countries, the use of animals for educational purposes in the veterinary curriculum requires approval by the Animal Ethics Committees. Many alternatives have been developed and are already in use in veterinary education such as interactive videos and computer simulations, in vitro cell cultures, slaughterhouse material, plastinated organs, dead animals from a humane/ethical source and clinical case-based practice.
The debate on the use of animals in veterinary education should include the question who benefits: the laboratory animals, the animal patients, the animal owners, the veterinary students, the teachers and/or the management?
Concerning the quality of the veterinarian trained with or without live animals, no difference in surgical performance could be shown so far.
Future aims should include a listing of case studies on available teaching materials in a database and a large-scale controlled international study on replacement methods in veterinary education.

Keywords: veterinary education, laboratory animals, alternatives

Anatomy is one of the oldest medical sciences. Stone-age people were already drawing animal forms about 30,000 years ago, which show certain knowledge of topographical anatomy. Aristotle (384-322 BC) performed detailed dissections on animals in order to extrapolate the results to humans, as in his time it was forbidden to carry out post mortems on humans. He described his findings in his books *Historia Animalium and De Partibus Animalium*. He was one of the first to use experimental animals.

Worldwide about 75 million vertebrates are used as experimental animals per year, 10 million of which within Europe. On average, about 2% is used for education and training purposes.

The basis of legislation on the use of animals for experimental purposes is the Three Rs principle of Russel and Burch:

replacement, reduction and refinement, launched in their book "The principles of humane experimental technique" in 1959.

Replacement means the substitution of live animals by *in vitro* techniques such as the use of cells and tissues, computerised models, videos, simulations and dummies.

The European legislation on experimental animals is based on two documents. In 1986, the Council of Europe launched the Convention for the Protection of Vertebrate Animals used for Experimental and other Scientific purposes (ETS123), followed by the EC Directive for the Protection of Vertebrate Animals used for Experimental and other Scientific purposes (86/609/EEC). These legislative regulations are based on the premise that under certain conditions it is morally acceptable to use animals for experimental and other scientific purposes.

Most laws contain provisions to protect the animals, such as the definition of legitimate purposes for animal use, competence of scientists and animal care staff, the use of alternatives and prevention of unnecessary pain and distress. One of the legitimate purposes is the use of animals in education and training (art. 2, ETS123). In biomedical education, a variety of disciplines use animals and animal tissues in order to acquire knowledge and develop skills. However, procedures involving animals carried out "for the purpose of education, training or further training for professions or other occupations are only permitted when the objective cannot be achieved by comparable effective audiovisual or any other suitable methods" (art. 25, ETS123).

In veterinary medicine the use of animals mainly focuses on learning and practising skills such as animal handling, animal behaviour, dissecting and surgical skills, but may also include understanding of anatomy, physiology, pharmacology and biochemistry. Eventually, the knowledge obtained in this way should benefit the animal itself in the long term rather than human beings, as is the case in the education of medical doctors.

The use of animals in veterinary education is becoming a subject of moral debate and is often opposed on educational and practical grounds. However, the discomfort that animals experience in relation to the purpose of their use should play a major role in this debate. For example, the grade of discomfort will be different for animals used to practice handling skills or for surgical training.

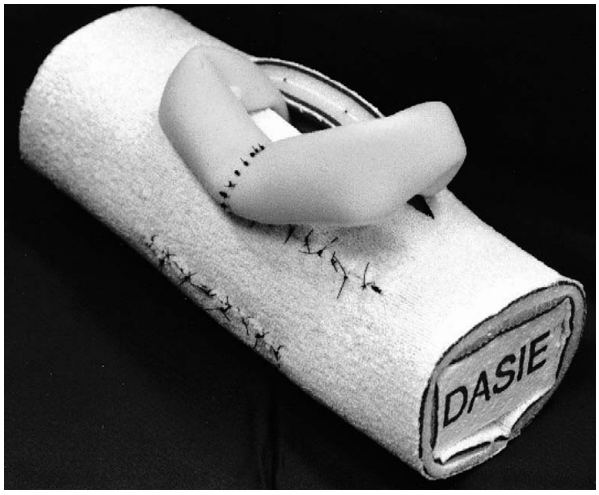


Fig. 1: DASIE™
(dr. David L. Holmberg e-mail: holmberg@uoguelph.ca)

Fig. 2: Critical Care Jerry™ (www.rescuecritters.com)

In many countries, the use of animals for educational purposes in the veterinary curriculum requires approval by the Animal Ethics Committees or Institutional Animal Care and Use Committees. The ethical admissibility of an educational program using animals is judged by balancing the purpose/relevance of the procedure against the suffering of the animal.

Many alternatives have been developed and are already in use in veterinary education such as interactive videos and computer simulations, *in vitro* cell cultures, slaughterhouse material, plastinated organs, dead animals from a humane/ethical source (body donation by owners or veterinary practices), animal shelters and clinical case-based practice, e.g. spaying/neutering of dogs and cats in animal shelters, patient-based practice at a later stage in the veterinary curriculum.

Dummies/simulators can be used to practice skills such as DASIE™ (Dog Abdominal Surrogate for Instructional Exercises) (fig. 1) to teach sterile techniques, surgical draping, instrument handling, incision making, suturing different layers or Critical Care Jerry™ (dog) (fig. 2) or Fluffy™ (cat) on which air and fluid aspiration from the thoracic cavity, jugular vascular access or intubation can be trained.

The rubber Koken™ rat (fig. 3) can be used to train handling, oral gavage, blood sampling from and *i.v.* injection into the lateral tail vein. Bicycle inner tubes covering foam are useful to practice suturing techniques (fig. 4). Both methods are in use in courses on laboratory animal science for scientists and animal care staff (FELASA, Cat. C, A, B).

The artificial rat model (IMTC, the Netherlands) is an aid in training microsurgical skills such as anastomoses and cannulation of the jugular vein.

However, would it be feasible and desirable to replace all experimental animal use in veterinary education? The debate on the use of animals in veterinary education should include the question of who benefits:

Fig. 3: Koken™ rat (www.kokenmpc.co.jp)

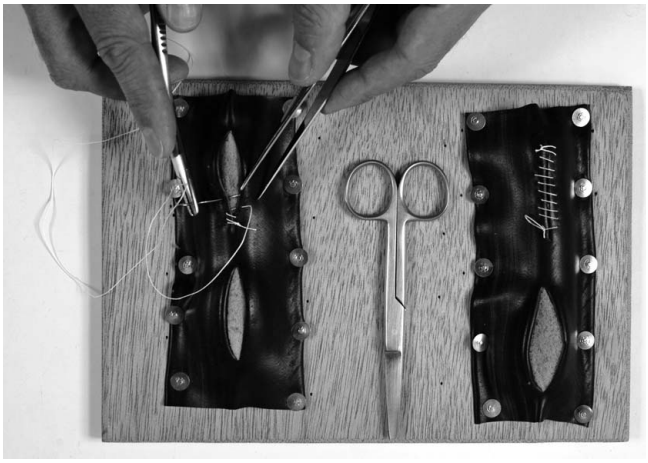


Fig. 4: Bicycle inner tube for suturing (Department Animals, Science & Society, University of Utrecht)

- The laboratory animals? Certainly they will benefit from alternatives as their lives are saved.
- The animal patients? Will they be treated and cared for in a better way when their veterinarians have been trained on live animals in their education?
- The animal owners? Will they accept mistakes in clinical case-based experience being part of the veterinary education, especially when it concerns their own animals?
- The veterinary students? Will the use of animals provide them with more skills and will they be better veterinarians?
- The teachers? Is it rewarding to teach without the use of animals, achieving the same educational results?
- Cost effective? As animals and their maintenance are expensive, the use of alternatives might be cost-effective.

Concerning the quality of the veterinarian trained with or without live animals, no difference in surgical performance could be shown (Balcombe, 2000; Griffon, 2000; Pavletic et al., 1994; White et al., 1992). Evaluation should be performed by a blind assessment by the Veterinary Faculty itself and/or on feedback of the employers' satisfaction in veterinary practices. Rating educational merits by the students themselves might be biased by the wish not to use animals.

Future aims should include a listing of case studies on available teaching materials in a database (EURCA, www.eurca.org) and a large-scale controlled international study on replacement methods vs. traditional animal-based teaching in veterinary education with a sufficient sample size, a blind study design involving several instructors and institutions and including short- and long-term evaluation.

If we could educate veterinary students as effectively without the use of animals, why should we not try?

References

Balcombe, J. P. (2000). Animals in Higher Education: Problems, Alternatives and Recommendations. Washington DC: Humane Society Press.

European Convention for the Protection of Vertebrate Animals used for Experimental and other Scientific Purposes (ETS 123) (1986). Strasbourg: Council of Europe.

European Council Directive (1986). Directive on the approximation of laws, regulations and administrative provisions of the member states regarding the Protection of vertebrate animals used for Experimental and other Scientific Purposes (86/609/EEC).

Griffon, D. J., Cronin, P., Kirby, B. and Cottrell, D. F. (2000). Evaluation of a hemostasis model for teaching ovariohysterectomy in veterinary surgery. *Vet Surg. Jul-Aug; 29(4)*, 309-16.

Pavletic, M. M., Schwartz, A., Berg, J. and Knapp, D. (1994). An assessment of the outcome of the alternative medical and surgical laboratory program at Tufts University. *J. Am. Vet. Med. Assoc. Jul 1;205(1)*, 97-100.

White, K. K., Wheaton, L. G. and Greene, S. A. (1992). Curriculum change related to live animal use. A four-year surgical curriculum. *J. Vet. Med. Educ. 19*, 6-10.

Correspondence to

Vera Baumans, DVM PhD DipECLAM
Professor in Laboratory Animal Science
Veterinary Resources
Karolinska Institutet
17177 Stockholm, Sweden
e-mail: vera.baumans@ki.se

Educating the Veterinary Professional about Animal Welfare*

Jasmijn de Boo[1] *and Andrew Knight*[2]
[1]World Society for the Protection of Animals (WSPA), London, UK, [2]Animal Consultants International, London, UK

Summary
For many years, animal welfare organisations have taken the lead in bringing about improvements in animal welfare, as well as helping to change attitudes towards animals. Unfortunately, the veterinary profession has often lagged behind in its support for reform. It is vital therefore that veterinary training includes a good grounding in animal welfare education. In 2003, the World Society for the Protection of Animals (WSPA) and the University of Bristol launched the 'Concepts in Animal Welfare Syllabus' to assist with the teaching of animal welfare in veterinary faculties. The syllabus stimulates focused critical thinking on animal welfare issues.

Keywords: animal welfare, education, veterinary profession

Introduction

Animal welfare has developed into a science in its own right, and, as a result, there is a growing amount of research into this subject. This research is funded by governments and other agencies and is often used as the basis for the reform of animal welfare legislation, and of the conditions under which animals are reared for food, used in scientific research, kept in captivity for entertainment or other purposes, or used as companion animals. As the body of welfare knowledge increases, its inclusion in educational curricula is likely to increase (Hewson, 2006). Welfare issues are also becoming more important considerations of many national and international veterinary associations.

Veterinarians are widely considered to be informed, rational authorities who possess expertise on virtually any topic relating to animals, including animal welfare. Career options for veterinarians include: working in veterinary practices, in research institutes, as welfare advisers for industry (e.g. feed or pharmaceutical) or for the government (e.g. in transport, slaughter and meat inspection). Many veterinarians are driven by a genuine desire to help animals, and this goal can be achieved with a proper knowledge of animal welfare.

Animal Welfare positions of Veterinary Associations

We reviewed the animal welfare positions of the World Veterinary Association (WVA), the American Veterinary Medical Association (AVMA), the British Veterinary Association (BVA), the Australian Veterinary Association (AVA) and the New Zealand Veterinary Association (NZVA). Many of the policies of the NZVA were in line with or exceeded the guidelines developed by that country's National Animal Welfare Advisory Committee (NAWAC), and we used the more-readily available NAWAC guidelines to represent the NZVA's positions in our survey. The five animal-use practices considered were: so-called "battery" cages for laying hens; gestation crates for pregnant sows; small crates and nutritionally-deficient diets for "veal" calves; the use of animals in scientific research and education; and the tail-docking of dogs.

We compared the results of the veterinary review to the available research data on public attitudes towards the five animal-use practices. A detailed report of the results is beyond the scope of this paper, but is available from the authors on request. Briefly, there was widespread and persistent public concern about many aspects of each of the five animal use practices, in all surveyed countries. In contrast, many of these specific concerns were not addressed clearly in the five veterinary associations' positions. All of the veterinary associations either lacked positions on or were not categorically opposed to the close confinement of laying hens, pregnant sows and "veal" calves, although the NZVA did recommend time limits on the use of sow gestation crates, and both the NZVA and the AVA recommended group, rather than individual, housing of "veal" calves. The only practice to which the public and the associations appeared to share a common opposition was the cosmetic tail docking of dogs, although the AVMA did not take a firm stance against this. In the case of animal experimentation, both the general public and the veterinary profession appear to support experimentation for human medical research to some degree, although public opinion remains very critical.

Our results clearly suggest that veterinarians lag behind the general public in their desire for animal welfare reform, unless the positions of veterinarians are not accurately represented by the veterinary associations surveyed. Anecdotal evidence and

*"Components of this paper were originally published in an article entitled "Concepts in Animal Welfare": A Syllabus in Animal Welfare Science and Ethics for Veterinary Schools" by Jasmijn de Boo and Andrew Knight in *J. Vet. Med. Educ. 2005 32(4)*, 451-453, and are reprinted with permission of the Association of American Veterinary Medical Colleges".

some research (Farm Sanctuary, 2004) indicate that a proportion of veterinarians do care deeply about these issues, and support animal welfare organisations, in some cases taking the lead on animal welfare issues. Furthermore, international veterinary bodies like the WVA, the Commonwealth Veterinary Association, the World Small Animal Veterinary Association and the Federation of European Companion Animal Veterinary Associations have all organised major animal welfare symposia in recent years. Thus, there may be some inconsistency between the apparent level of concern about animal welfare expressed by the surveyed veterinary associations and the level of concern of individual veterinarians and some other veterinary associations. Nevertheless, even if the surveyed associations do not fully represent the positions of the majority of veterinarians, it appears likely that a substantial proportion of veterinarians lags behind the general public in their desire for the advancement of animal welfare issues. We attribute this in part to inadequate education about the science and issues of animal welfare during formal veterinary training.

Trends in veterinary education

The veterinary profession has its origins in agricultural practice, although in the developed world today most veterinarians work primarily with companion animals. Veterinary medical curricula have been modified accordingly over time. Accompanying these changes is the marked feminisation of a previously male-dominated profession. This gender shift is bringing some changes in the attitudes of veterinary students towards animal welfare. For example, a cross-sectional study of veterinary students in their first preclinical year, first clinical year and final year of study showed that the women in each of these groups rated themselves as having significantly higher levels of emotional empathy with animals than did the men. This difference was most marked in the final-year students; moreover, the males in that group showed lower levels of empathy than their peers in the earlier year-groups (Paul and Podberscek, 2000). Research at a US veterinary school has examined veterinary students' attitudes to pain management. Fourth year students were less likely than second or third year students to provide analgesia for certain surgeries (Hellyer et al., 1999). Moreover, it appears likely that there is inhibition of the normal development of moral reasoning ability during the four years of veterinary school (Self et al., 1991).

These findings may be attributable in part to the attitudes of teaching faculty and the examples they set. The apparently reduced concerns for animal welfare might also, in some cases, represent adaptations that enable veterinary students to withstand what could otherwise be intolerable psychological stresses that result from being required to harm sentient creatures in the absence of overwhelming necessity (Capaldo, 2004).

During their training, veterinary students are frequently required to harm and kill animals in preclinical subjects such as anatomy (dissection, often of purpose-killed animals or animals from ethically-questionable sources), physiology, biochemistry and pharmacology ("demonstration" experiments on living animals, usually of long-established scientific concepts, with animals usually killed during or at the end of the experiment). Students have also traditionally been required to practice clinical, surgical and anaesthetic skills by anaesthetising healthy animals, conducting surgical procedures on them, and killing any survivors at the end (not all survive these frequently-lengthy operations) (Knight, 1999).

While many veterinary schools are continuing to refine their curricula to avoid harmful animal use, such use may still be found in veterinary education worldwide, both in preclinical and clinical (surgical) training. Furthermore, students who are not aware of the existence of alternative methods to harmful animal use in education and scientific research are less likely to consider the 3Rs when planning research projects themselves at graduate and postgraduate level.

Discussion

Where harmful animal use is retained in the curriculum and animal welfare education is lacking, it is likely that graduating veterinarians will have a diminished appreciation of animal sentience, and a diminished understanding of animal welfare science and animal welfare issues, all of which will impede their knowledge of the 3Rs and their abilities to guide their clients and the wider public appropriately. This may explain why the veterinary associations we surveyed seemed to lag behind the public in their concern about the welfare of animals in several management systems widely believed to result in poor welfare.

Recommendations

Although animal welfare is necessary as part of formal veterinary education throughout Europe as part of the European Community's move towards harmonisation of professional qualifications, animal welfare education is underdeveloped in most veterinary schools and, we believe, has not received the attention it deserves in the curriculum. However, there are increasing numbers of courses on animal welfare being implemented around the world. Some of these courses are integrated into undergraduate veterinary education. In addition, there are postgraduate courses such as the MSc in Applied Animal Behaviour and Animal Welfare at the University of Edinburgh, Scotland, and the MSc in Animals and Public Policy at the Cummings School of Veterinary Medicine at Tufts University, USA. Moreover, one of the animal welfare mandates of the World Animal Health Organisation (OIE) is promotion of the inclusion of animal welfare in undergraduate and post-graduate veterinary curricula (Anon, 2005).

To encourage the introduction of animal welfare education into veterinary curricula worldwide, the World Society for the Protection of Animals (WSPA) developed an outline of the "Concepts in Animal Welfare" syllabus in 2000. The complete syllabus was developed in collaboration with the University of

Bristol, School of Veterinary Science, and was launched on CD ROM in 2003. The aims of the CD ROM are to help students:

- to develop an understanding of animal welfare relevant to an animal's physiological and psychological well-being;
- to recognise the welfare, ethical and legal implications of animal use practices and to be able to apply critical analyses from each perspective, for different species in different situations; and
- to stimulate focused critical thinking on welfare issues, which can be developed throughout the course and the individual's professional career.

The syllabus comprises 30 theoretical teaching units in PowerPoint format, illustrated topics with practical examples and case studies, questions and assessment materials, suggested reading lists, and additional relevant materials. There are seven core and 23 elective teaching modules covering a wide range of animal welfare topics (tab. 1).

The materials are suitable for use in class as well as for independent study, and they also stimulate interaction between students and with the lecturer. Lecturers can easily adapt the materials to suit their needs, by adding relevant country-specific information or by omitting modules for which there is insufficient time. The CD ROM can be used both at undergraduate and postgraduate level, and it is suitable for use on its own or for integration into existing courses on behaviour, physiology, ethics, or veterinary law. Navigation of the CD ROM is easy, and the presentation of the materials is engaging. The modules build on the experiences of both the University of Bristol, which is one of the world's leading centres for the study of animal welfare science, and WSPA, which has 50 years of international experience in

Tab. 1: Modules of the 'Concepts in Animal Welfare' syllabus

Module number	Topics	Core module
1	Animal welfare introduction	Yes
2	Welfare assessment and the Five Freedoms	Yes
3	Physiological indicators of welfare (1)	No
4	Physiological indicators of welfare (2)	No
5	Immune and production indicators of welfare	No
6	Behavioural indicators (1)	No
7	Behavioural indicators (2)	No
8	Group assessment and management of welfare	No
9	Human-animal interactions	Yes
10	Introduction to animal welfare ethics	Yes
11	Interaction with other ethical concerns	No
12	The role of the veterinary profession and individual veterinarian	Yes
13	Humane education	No
14	Animal welfare organisations	No
15	Protection legislation (1)	Yes
16	Protection legislation (2)	Yes
	-Enforcement and political pressure	
17	Commercial exploitation of wildlife	No
18	Influence of the marketplace	No
19	Farm animal welfare assessment and issues (1)	No
20	Farm animal welfare assessment and issues (2)	No
21	Farm animal transport and markets	No
22	Slaughter of farm animals	No
23	Working animals	No
24	Animals used in entertainment	No
25	Animals in experiments	No
26	Companion animals (1)	No
	-Population control programs	
27	Companion animals (2)	No
	-Wider considerations	
28	Euthanasia	No
29	Wild animal management	No
30	a) Religion and animals	No
	b) War and natural disasters	

advancing animal welfare issues, including collaboration with governments, international organisations and veterinary professional bodies.

To facilitate the implementation of the syllabus, WSPA has organised a series of conferences and workshops since 2004 for over 300 veterinary faculties in Brazil, Central and South American countries, the Czech Republic and other Eastern European countries, Indonesia, India and the Philippines. As a result of the workshop in the Philippines in March 2005, a steering committee was formed, which is currently reviewing the national Philippine veterinary curriculum with a view to including more animal welfare aspects in existing modules on husbandry and breeding. Further conferences and workshops are planned for 2006 in Japan, Brazil, Latin American and South American countries, and Africa. In Colombia, animal welfare will be a compulsory part of the curriculum from 2006.

WSPA does not yet have a complete overview of which universities have implemented all or parts of the syllabus, but we know of many universities, in Australia, Brazil, Canada, Colombia, India, Indonesia, Kenya, New Zealand, the Philippines, South Africa, the United States of America, and many European countries, that are successfully using the resource. An extensive assessment form for users has been developed to help WSPA improve the syllabus even further.

Conclusion

Informed positions on questions of animal use in education and scientific research are associated with humane attitudes and a sound knowledge of animal welfare science. Although that discipline is well-established, with an ever-expanding body of related research, most veterinarians will remain relatively ignorant of animal welfare science and issues unless they learn about them during their formal education. We hope that the "Concepts in Animal Welfare" syllabus will play an important role in assisting veterinarians to develop a sound understanding of this increasingly important field. The syllabus is designed to achieve this by stimulating students to undertake focused critical thinking on welfare issues, not only during their veterinary course, but throughout their entire career.

References

Anon. (2005). *Report of the Third Meeting of the OIE Working Group on Animal Welfare*. Extract from the report of the January 2005 meeting of the OIE Terrestrial Animal Health Standards Commission. Paris. Retrieved 12 October 2005 http://www.oie.int/eng/bien_etre/AW_WG_december2004_eng.pdf.

Capaldo, T. (2004). The psychological effects on students of using animals in ways that they see as ethically, morally or religiously wrong. *ATLA, Suppl 1*, 525-531.

Farm Sanctuary (2004). *Opinions of Veterinarians and Positions of the AVMA: Analysis of Eight Commonly Occurring Farming Practices*. Watkins Glenn, NY, US: Farm Sanctuary.

Hellyer, P., Frederick, C., Lacy, M. et al. (1999). Attitudes of veterinary medical students, house officers, clinical faculty, and staff toward pain management in animals. *J. Am. Vet. Med. Assoc. 214(2)*, 238-244.

Hewson, C. (2006). Why the Theme Animal Welfare. *J. of Veterinary Medical Education 32(4)*, 416-418.

Knight, A. (1999). Alternatives to harmful animal use in tertiary education. *ATLA 27(6)*, 967-974.

Paul, E. and Podberscek, A. (2000). Veterinary education and students' attitudes towards animal welfare. *Vet. Rec. 146(10)*, 269-72.

Self, D., Schrader, D., Baldwin, S. et al. (1991). Study of the influence of veterinary medical education on the moral development of veterinary students. *J. Am. Vet. Med. Assoc. 198(5)*, 782-787.

Correspondence to

Jasmijn de Boo, BSc. MSc. DipEd.
World Society for the Protection of Animals (WSPA)
14th Floor 89 Albert Embankment
London SE1 7TP, UK
e-mail: jasmijndeboo@wspa.org.uk

User-friendly Curricula on Alternatives for Research Scientists

Lynette A. Hart, Mary W. Wood and Hsin-Yi Weng
UC Center for Animal Alternatives, School of Veterinary Medicine, University of California, Davis, USA

Summary
Engaging research scientists in the field of animal alternatives requires relevance and timing that complement the research efforts of each particular person. Presentations are effective when convenient and in a context valued by scientists. Targeted curricula can be designed as part of courses, workshops, or symposia on topics of interest. Websites accompanying the material make it feasible for participants to retrieve the information when needed for an animal use protocol. For locally relevant material, scientists partnered with librarians can build resources that piggyback on existing tools, adapting them for the research interests of participants and their contexts.

Keywords: alternatives, curricula, presentations, scientists, workshops, courses, symposia, 3Rs, research

Introduction

Clarifying the United States Animal Welfare Act (2002), Policies 11 and 12 (United States Department of Agriculture, 1997, 2000) require that researchers conduct specific searches for alternatives to procedures that may cause more than momentary pain or distress. This requirement assigns researchers a searching task on a somewhat different topic than they would normally follow in the literature and is sometimes viewed as a burden. In view of these requirements for compliance with the Animal Welfare Act, it is in the interests of academic institutions to facilitate this searching endeavour and enhance the efficiency of this process.

Several challenges face anyone wanting to advance curricular development for research scientists in the area of alternatives. Scientists work in highly individualised fields of research, using a particular sequence of methods and techniques. They may view searching for alternatives, at least initially, as an irrelevant obligation. But they must address alternatives when they are writing up an animal care and use protocol, so the timing of presenting information to them or building in an easy access for later when needed is critical. A potential for engaging scientists and increasing their interest in alternatives exists if the presentation involves improving access to the research literature in general, or addressing particular questions that matter to the scientist.

Alternatives have the potential to arouse interest and involvement when they are presented within the audience's context of research interests. Targeted presentations designed and presented on-campus for specific courses, workshops, or symposia can complement the other information being presented in that venue and profile the practical value of bibliographic searching techniques for accessing information on alternatives. In this paper we identify three challenges in disseminating alternatives and then present approaches we use at various venues for developing targeted curricula that complement specific topics being addressed for the particular audience. Finally, we propose strategies for generating local support and opportunities to present information on alternatives.

Three challenges in disseminating alternatives

To be effective, disseminating information on alternatives requires a type of teaching, yet there is not a ready-made audience. People at research institutions seeking to upgrade their institutional compliance with the requirement of searching for alternatives might first consider how best to recruit and stimulate a talented teacher and a willing audience for alternatives. In this light, we draw attention to three essential components of an effective program for enhanced searching for alternatives.

Researchers

The fact that searching for alternatives is a requirement sometimes positions this obligation as an onerous task. Scientists in universities who are required to complete animal use protocols are highly accomplished and knowledgeable. They are self-starters, highly focused and motivated to work intensely in their subject of interest. They are accustomed to setting their own objectives. They are proficient in tracking the literature and new developments in the areas relating to their science, and know how to quickly monitor and locate that literature. Their colleagues review their work prior to its publication. Extensive reviews also are strict prerequisites for funding and grant support. Only work that is well-regarded makes the cuts for funding and publication. After surviving such extensive peer-scrutiny, researchers may not readily welcome another requirement that to them may appear frivolous and not directly related to their research. Thus, when crafting curricula on alternatives for research scientists, it is useful to understand that they can be a

challenging audience to engage, or even to get to show up for presentations. The material must be targeted to their interests, and presented at a convenient venue, preferably sandwiched in at an occasion that already promises to attract their attendance.

Technology

The searching for alternatives is less straightforward than a typical search on a topic of someone's line of research. Topics on alternatives are dispersed across many interdisciplinary fields and covered in several general databases as well as some specialised ones (Hart and Wood, 2000). Conducting an effective search that is relevant to the species and topic of the study requires some thoughtful effort (Hart et al., 2005). It is not a reasonable expectation, even for a brilliant scientist, that someone could perform an effective search on alternatives *de novo*.

Nonetheless, the technology lends itself to creating efficient tools, such as search templates, filters, and grids (Wood and Hart, 2004a) and database grids (Hart et al., 2005a). These can simplify searching for alternatives, lead to more effective searches, and create a spillover of more effective searching in the scientists' areas of study.

Librarians as educators to assist in searching for alternatives

Given that searching for alternatives requires an unusual specialised expertise, most institutions lack specialists in this area. For institutions seeking to create pro-active local support toward more effective and efficient searching for alternatives, librarians have the requisite skills to quickly acquire some proficiency and offer assistance (Hart and Wood, 2000). Given the willingness of librarians to provide service and master new subject areas, it is a feasible objective for an administrator or animal care committee to recruit a librarian to become a campus specialist to provide assistance and instruction in the area of alternatives.

Methods and avenues of presentation

The general goal of offering demonstrations on searching for alternatives is to increase scientific literacy. The instruction is provided in the context of the requirement to search for alternatives. Typically, individuals think about searching for alternatives when facing the requirement and at other times the topic is not a high priority. One objective, then, is to have information accessible to the person who needs it, at the precise time when it is needed. The researcher needs to know about alternatives resources, know where to find them, and how to use them. Delivery on the web of the resources is ideally suited for offering access at any time from any place. A second objective during presentations is to increase the valence of the information, and demonstrate its practical value and relationship to science in general and its potential for assisting the person in being more effective in searching.

Due to the requirement to search for alternatives when preparing a protocol, there are some teachable moments when researchers are seeking to learn. Nonetheless, this topic does not promise to draw an enthusiastic audience. Building a constituency of interested people can be the biggest challenge. With this in mind, in each of the contexts for presentation described below, an effort is made to tailor the information to the occasion and audience, and to provide a web page on the specific topic. The web page allows easy review of the presentation and facilitates the person returning to the specific techniques when needed later. With the accessibility of the internet, the information can be delivered almost anywhere, anytime, as needed by any user.

Courses

Information on alternatives can be offered in courses for undergraduates, graduate students, and veterinary or other professional students. For example, the presentation for an undergraduate course on companion animal biology, ANS 42, Introductory Companion Animal Biology, focuses on companion animal welfare (Wood et al., 2005a). This designed website is used to demonstrate topics related to husbandry of dogs and cats, behavioural indicators of welfare, and behavioural effects of spay/neuter surgery.

Separate courses for veterinary and graduate students address mouse behaviour and biology: PHR 408, Mouse Behavior and Biology, and PMI 280, The Mouse as an Experimental Model for Human and Animal Diseases (Wood and Hart, 2005a). The presentation and specialised website on alternatives includes provides ready access to a variety of searching tools on mice and alternatives that we have developed, along with links to classic resources offered by the Jackson Laboratories and a variety of other specialised resources on mice.

A mentored clinical research training program associated with a masters degree comprises junior clinical faculty of the medical and veterinary schools who are being groomed to conduct their own research programs. A presentation to this group focuses on the animal care and use protocol and standard operating procedures used by various campuses (Wood and Hart, 2005b).

Workshops

Workshops can either be sponsored on the specific topic of alternatives, or presented as part of a broader topic for the animal care, primate, or general research community. Each year, we offer workshops for the veterinary laboratory animal residents from the California National Primate Research Center and from the Center for Laboratory Animal Science at the University of California, Davis (Wood and Hart, 2005b). These workshops provide a general overview of alternatives, animal subject protocols, and standard operating procedures, and offer hands-on instruction with computer searching. Most valuable, perhaps, is that participants can raise unique problems they are experiencing, that would not be covered in a general overview, e.g., dealing with ants in the animal room.

Additional workshops are developed and presented on-site with hands-on instruction as tailored tutorials for visiting veterinarians, including the USDA Animal and Plant Health Inspection Service Preceptor Veterinary Fellows each year, and visiting librarians. Presentations on other campuses with animal care personnel and librarians are tailored to the particular needs and research interests of the community.

Symposia

Symposia or conferences can be effective as collaborative ventures focusing on timely topics. We use this approach for an annual symposium co-sponsored with the California National Primate Research Center. Topics are selected on emerging new techniques that are relevant to alternatives, and for which we develop tailored presentations on, for example, new methods of imaging (Wood and Hart, 2004b), or cell culture and explants (Wood et al., 2005b). Once again, a special web page is prepared and configured that users can conveniently access following the presentation. This method of instruction offers support to users in efficient searching within their context of the course material or working setting, such that the alternatives curricula supplement their needs.

Conclusions: Generating local support and opportunities to present information

The sequence of providing effective, user-friendly curricula on alternatives for research scientists gets underway by identifying a librarian who would be interested in learning and educating about alternatives. A second step is identifying opportunities for presenting material to interested audiences of researchers. Personalised presentations as opportunities arise for campus courses, and workshops or symposia on special themes, can gradually build an interested constituency that recognises and appreciates the value of the information on alternatives.

References

Hart, L. A. and Wood, M. W. (2000). Facilitating compliance in bibliographic searching for alternatives: A model program for developing local support. In M. Balls, A.-M. Van Zeller and M. E. Halder (eds.), *Progress in the Reduction, Refinement and Replacement of Animal Experimentation* (1409-1414). Amsterdam: Elsevier.

Hart, L. A., Wood, M. W. and Weng, H.-Y. (2005a). Effective searching of the scientific literature for alternatives: search grids for appropriate databases. *Anim. Welf.*, in press.

United States Animal Welfare Act. 7USC 2131-2159, adopted 1966, amended 2002. Available at: http://www.aphis.usda.gov/ac/awa.html.

United States Department of Agriculture. (1997). *Animal and Plant Health Inspection Service Animal Care, USDA Animal and Plant Health Inspection Animal Care Policy Manual Policy #11–Painful Procedures*–April 14, 1997. Website accessed 08/12/2005: http://www.aphis.usda.gov/ac/policy/policy11.pdf.

United States Department of Agriculture. (2000). *Animal and Plant Health Inspection Service Animal Care, USDA Animal and Plant Health Inspection Animal Care Policy Manual Policy #12–Consideration of Alternatives to Painful/Distressful Procedures*–June 21, 2000. Website accessed 08/12/2005: http://www.aphis.usda.gov/ac/policy/policy12.pdf.

Wood, M. W. and Hart, L. A. (2004a). Search filters, a new tool in the search for alternatives: Locating mouse strains as disease models. In A. Rowan and H. Spielmann (eds.), *Alternatives and Animal Use in the Life Sciences* (599-602). Amsterdam: Elsevier.

Wood, M. W. and Hart, L. A. (2004b). Noninvasive imaging: practical approaches for the 3Rs. Website accessed 08/15/2005: http://www.vetmed.ucdavis.edu/Animal_Alternatives/imaging.html.

Wood, M. W. and Hart, L. A. (2005a). PHR 408. Behavior, biology and welfare of mice: the laboratory animal of the future. Website accessed 08/15/2005: http://www.vetmed.ucdavis.edu/Animal_Alternatives/phr408-Mice.html.

Wood, M. W. and Hart, L. A. (2005b). UC research, protocols and the 3Rs alternatives search resources. Website accessed 08/15/2005: http://www.vetmed.ucdavis.edu/Animal_ Alternatives/ucprotocols.html.

Wood, M. W., Weng, H.-Y., Bain, M. J. and Hart, L. A. (2005a). Welfare of dogs and cats. Website accessed 08.17-2005: http://www.vetmed.ucdavis.edu/Animal_Alternatives/welfare.htm.

Wood, M. W., Weng, H.-Y. and Hart, L. A. (2005b). Cell culture. Website accessed 08/15/2005. http://www.vetmed.ucdavis.edu/Animal_Alternatives/cell.htm.

Acknowledgements

We are indebted to our colleagues at the University of California who have supported our efforts toward increasing the visibility of instructional offerings for more effective bibliographic searching for alternatives. Librarians at each University of California campus, as well as the animal care community that participates in animal subject protocol review, have collaborated in enhancing access to information on animal alternatives. Partial financial support for these efforts was provided by the McBeth Foundation, Bayer Corporation, and Animalearn. Collaborators in workshops included the Scientists Center for Animal Welfare, the California National Primate Research Center, and the Animal Welfare Information Center.

Correspondence to

Lynette Hart, PhD
UC Center for Animal Alternatives
School of Veterinary Medicine
University of California
Davis, CA 95616, USA
e-mail: lahart@ucdavis.edu

Animals and Alternatives in Biomedical Education in the Baltics

Osvaldas Ruksenas
Department of Biochemistry-Biophysics, Vilnius University, Vilnius, Lithuania

Summary
The use of animals in biomedical education has decreased substantially during the last decade in Baltic countries. There are several reasons for this decrease: i) introduction of legislation regulating use and care of laboratory animals; ii) demand from students for the use of alternatives in teaching; iii) improvement of alternative methods and equipment; iv) feasibility of alternative teaching materials. Relatively high price of alternatives, lack of computers and modern A/V equipment at universities, difficulties of students with foreign languages are among the main factors limiting the increase in the use of alternatives.

Keywords: laboratory animals, alternatives, education

Introduction

Three countries – Estonia, Latvia and Lithuania – located along the East cost of the Baltic Sea share many common features in history, economy, and culture therefore quite often these countries are taken as an unit and called Baltic countries or Baltics instead of listing all three names. Laboratory animals have been used in research and education in the Baltics for a long time but there was no laboratory animal science at that time. Whereas well-known political and social changes in the 90ies strongly influenced the situation with laboratory animals – people became aware that a special scientific field such as laboratory animal science does exist and this area can not be neglected. During the last 15 years laboratory animal science was established and developed in Baltic countries – there was established Baltic Laboratory Animal Science Association (Balt-LASA), Lithuanian Laboratory Animal Science Association (Lith-LASA), the first textbook "Basics of laboratory animal science" was published in Lithuanian language, and a legal basis regulating use and maintenance of laboratory animals has been created (Ruksenas and Simkeviciene, 2003).

As long as laboratory animals are used in biomedical education the special emphasis on implementation of the principle of the 3Rs – replacement, reduction and refinement – has to be done. In this aspect changes in the field of laboratory animal science occurred in Baltic countries and are of high value.

Animals in biomedical education

Two main forms of teaching are used in biomedical sciences – lectures and practicals. With respect to the use of animals lectures are less interesting, because irrespective of the style of lectures – classical or modern – based on the latest audio/video tools, lectures usually do not accommodate use of live animals. Therefore the practicals are the main activity related to the use of laboratory animals.

From the point of view of methods used practicals can be subdivided into two opposite groups – based only on the use of animals or those using alternative methods, with the broad spectrum of practicals using both – animals and alternatives in between.

When we consider the use of animals in practicals few questions raise up: i) is it justifiable to use animals for this purpose? ii) what are the advantages and disadvantages of using animals?

Considering the first question there are very active and long-lasting discussions with solid arguments on both sides. As a result it is difficult to give a finite response to this question. Leaving aside ethical, moral issues one of the main questions helping to solve this problem could be – how does the use of animals in teaching comply with the objectives of teaching? After analysis of many curricula for biological and medical programs the clear answer emerges that in many cases objectives of programs could be achieved equally well without using animals. This is a serious argument for the reduction of the number of animals used in biomedical teaching.

However, for time being animals are not taken out completely from the teaching process and this poses a question about the advantages and disadvantages of this way of teaching over alternatives. The main and probably the only one advantage of using animals is "realism". This means that in performing animal-based experiments students are able to experience what a live animal is – what is it's body composition, temperature, heart beat, how does it breaths etc. In addition to this students can learn (at least to some extent) to handle animals and this in turn can change the attitude towards the animals.

Naturally, the use of animals has a number of disadvantages as well and finances are among the most important ones. In 2004 all three Baltic countries joined the European Union (EU) and this implies that these countries have to follow European legislation, including legal acts regulating the use and care of animals used in experiments. Taking into account the fact that just a little more than 10 years ago there were no regulations concerning laboratory animal science a substantial investment is required into the area of laboratory animals in order to conform to these regulations. Even if there is compliance with EU regulation, animals are "costly" experimental objects, because they are expensive themselves, maintenance is costly etc.

Another factor against the use of animals in education is a growing awareness of the society and its concern about animals. There are no extreme acts or manifestations against the animal users in the Baltics, as they are in some other European countries. But this issue is discussed in the media from time to time therefore it has to be taken into account. Finally, there is one more disadvantage of animal use in teaching – this is the well-known variability in biological data and to some extent the unpredictability of experimental outcome. Usually students have only limited time for practicals, whose results they have to report. On the other hand most of them are not experienced in animal experimentation. Altogether this leads to the possibility to fail during practicals which could have distressing effect on students.

Despite the discussed advantages and disadvantages, animals continue to be used in biomedical education and this raises another question – how many animals are used for this purpose, what species of animals are used? To answer this question some surveys have been performed. In 1996 the Lithuanian Laboratory Animal Science Association (Lith-LASA) performed a survey aiming to evaluate what number of laboratory animals and for what purposes they were used in Lithuania. It turned out that 21,000 of animals were used in total and 2,300 or 11% of these were used for teaching (Simkeviciene et. al., 1998). Figure 1 illustrates the distribution of the animals used by species. This distribution is rather typical – rodents are dominating over other species (comprise over 70% of all animals).

Another survey aimed to assess the use of laboratory animals in teaching has been performed in 2004. The numbers of animals were much lower if compared to the year 1996 – 8,000 animals in total and 1,360 or 17% were used for teaching. Figure 2 illustrates distribution of animals used for teaching by species. Distribution is very uneven – frogs comprise 80% of all animals used in teaching.

The comparison of data from 1996 and 2004 leads to a very clear conclusion – the number of animals used in total and in teaching particularly, were substantially reduced during the eight year period. However, the extent of this reduction is different – the total number of laboratory animals used declined much more – by 64% as compared to the reduction of number of animals used in teaching (by 41%). The reason for this difference goes back to the 90ies, when well-known political and economical changes in all three Baltic countries occurred. Among the numerous consequences of these events were serious changes in the area of science and education. In biomedical sciences this was expressed by the reduction of experiments performed using animals. As a result less animals were used. The coverage of education did not change so drastically therefore the reduction of the number of animals used in teaching is less pronounced.

The situation in Latvia and Estonia is similar to the one described in Lithuania, but the numbers of animals used for teaching are different: over 100 animals/year were used in Latvia and no animals were used in biomedical education in Estonia. The much lower number of animals used for teaching in Latvia can be explained by the smaller number of educational institutions delivering biomedical programs. The use of animals in biomedical education was stopped completely a few years ago in Estonia. Since that time these animals are only used for purposes other than teaching in this country.

Overall it can be concluded that the number of animals used in education is reduced – what are the reasons for that?

The introduction of legislation complying with EU legislation regulating the use and care of laboratory animals is among the main factors causing a reduction in the number of animals used. For example, there are the following acts of law related to laboratory animals in Lithuania: Law on Animals Care, Handling and Use, 1997; Veterinary Regulations on Breeding, Handling and Transportation of Laboratory Animals, 1998; Rules of Good Laboratory Practice, 1999; Rules on the Use of Laboratory Animals in Scientific Experiments, 1999. These legal acts impose requirements for both care and maintenance of laboratory animals and competence of personnel working with animals. To comply with legal requirements financial resources are required and this is the second limiting factor.

A rather new and unexpected factor in the Baltics is a changing attitude of students towards the use of animals. Students started to question the sense of using animals in practicals, possibilities to replace by alternatives etc. As a result in some Universities the use of animals for teaching has been stopped. Two more reasons are

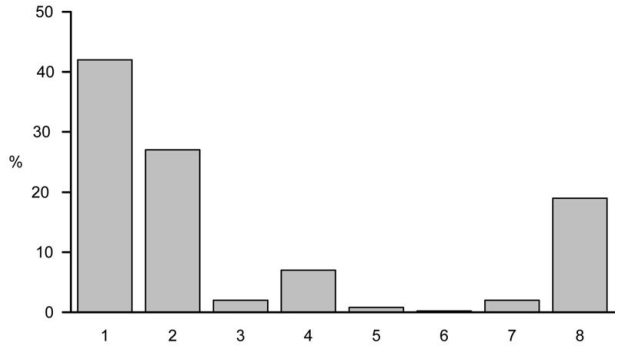


Fig. 1: Distribution by species of laboratory animals used in Lithuania in the year 1996. 1 – mice, 2 – rats, 3 – guinea pigs, 4 – rabbits, 5 – dogs, 6 – cats, 7 – birds, 8 – others.

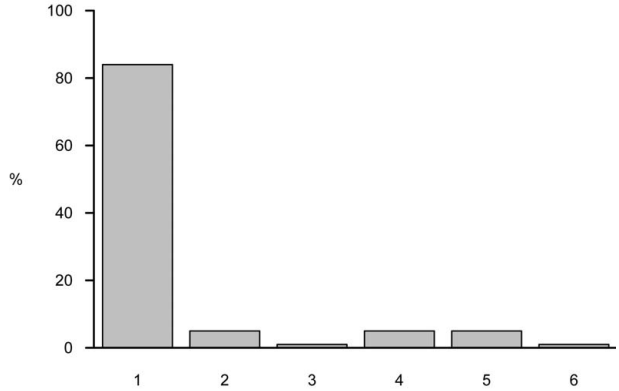


Fig. 2: Distribution by species of laboratory animals used for education in Lithuania in the year 2004. 1 – frogs, 2 – fishes, 3 – rats, 4 – mice, 5 – birds, 6 – rabbits.

related to alternatives: i) constantly increasing quality of alternatives makes them more and more suitable for replacement of animals; ii) the growing economy of Baltic countries and the reduction of prices of alternatives increase the feasibility of alternatives which in turn causes a reduction of the number of animals used.

Alternatives in biomedical education

Alternatives used in biomedical education in the Baltics can be grouped into three categories – computer-based simulations and self-experiments, models/mannequins and cell cultures. The most popular are computer-based alternatives and models, whereas cell cultures just started to be used. What are the advantages of using alternatives in education? First of all, it is the attractiveness of these methods, this especially concerns computer-based alternatives. This is defined by technological changes – the young generation is growing up in a computerised environment – already from kindergarden they are used to computers, the amount of information available via computers is growing every day. Therefore students easily and enthusiastically accept information provided via computer-based setups. As a result this increases their interest in performing practicals and finally taking in information. Flexibility is another factor contributing to increasing popularity of computer-based alternatives. Flexibility means that having even a basic setup for self-experimentation it is easy to configure it for performance of different experiments in a short time.

And last but not least – the economical factor. If animals are used in teaching they have to be kept in proper conditions, i.e. it is necessary to have well equipped animal facilities run by qualified specialists. And both – facility and specialists are very costly compared to alternatives, especially in the long perspective. Taking into account permanent shortage of finances at universities this is serious argument in favour of alternatives.

Alternatives are attractive, but they are not ideal, so what are the disadvantages of using alternatives? The first and main is what could be called "lack of life". This means that even the most sophisticated alternative can not simulate the feeling one experiences by handling a live animal. For example any of the best virtual models of rat or dog used for practicing purposes are far behind live animals in this respect. This aspect is of particular value for students planning to work with animals in their future. Another disadvantage could be described as the ratio between cost and quality of alternatives. The problem is that there are rather many alternatives available on the market and new ones are constantly appearing, most of them are not cheap, but quality is not always corresponding to price. This problem becomes much sharper in the case of limited financial resources. And finally there is one more problem related to alternatives – language. Usually alternatives are developed in the main European languages which still are literally "foreign" languages for some part of our students because their knowledge of foreign languages is far from sufficient. This restricts the use of computer-based alternatives, because students can not understand comments or texts presented and as a result they can not work independently, without the permanent guidance/translation of a teacher. Hopefully this is a short-lasting problem which fades out in the next few years.

Finally, there is one more factor limiting the use of computer-based alternatives – in general, institutions of higher education in the Baltics are lacking computers and modern audio/video equipment in lecture halls and laboratories. But the recent development of the economy allows to expect that in the close future this obstacle will disappear.

Overall, advantages of alternatives are taking over disadvantages and as a result they are increasingly used in the universities of Baltic countries. The most popular and successfully used computer-based alternatives are the following: "How your body works", "The dynamic human", "The virtual physiology lab", "SimNerv", "Human Physiology", "Laboratory simulation for Human Physiology Exercises".

Turning to commercial equipment used for self-experimentation the list is much shorter – the setup from "BIOPAC" is the most popular in Estonia and Lithuania. There is one more custom developed setup "ComLab" available in Lithuania. This setup was developed within the Leonardo da Vinci program a project financed in cooperation with numerous European partners. It turned out to be a very successful and efficient product compared to commercial setups by technical characteristics, at a fraction of their costs.

Conclusions

1. The number of laboratory animals used for teaching in Baltic countries is constantly decreasing.
2. Alternatives are successfully implemented in curricula of biomedical programs.
3. The principle of 3R's is being successfully implemented in Baltic countries.

Acknowledgements

Many thanks to Prof. V.Klusa and Prof. A.Lang for providing data on numbers of laboratory animals used in education in Latvia and Estonia.

References

Ruksenas O. and Simkeviciene V. (2003). Development of laboratory animal science in Lithuania. *Scandinavian Journal of Laboratory Animal Science 30 (3)*, 157-164.

Simkeviciene, V., Laukaitis V. and Uleckiene S. (1998). Use of laboratory animals in biomedical research. *Medicina 34* (in Lithuanian), 795-799.

Correspondence to

Prof. Dr. Osvaldas Ruksenas
Department of Biochemistry-Biophysics
Vilnius University
Ciurlionio 21/27
Vilnius, LT-03101
Lithuania
e-mail: osvaldas.ruksenas@gf.vu.lt

Educating Scientists on Alternatives. A Continuous Process

Jan van der Valk
NCA, Dpt. Animals, Science & Society, Faculty of Veterinary Medicine, Utrecht University, The Netherlands

Summary
Already from the start of their study in one of the biomedical sciences should students be made aware of the fact that animal experimentation is not always an obvious choice in biomedical research and that in many instances 3Rs models can and should be used. During laboratory classes, whenever possible, animal free teaching models should be considered and used. Also, when it is regarded essential to have the students work with experimental animals, they should be informed on the consequences and the concerns of the society.
Every scientist designing animal experiments should be qualified and aware of the 3Rs. The Laboratory Animal Science course, where Replacement, Reduction and Refinement (3Rs) are the main themes, offers this qualification.
Scientist should be continuously stimulated to consider the 3Rs when planning animal experiments. The animal ethics committee requests that 3Rs models have been considered before an animal experiment proposal is approved. This requirement and subsequent verification of an expert on alternatives ensures that scientists are (made) aware of possible 3Rs models in their field of interest.
Several journals now require a statement that the 3Rs have been considered and applied. Furthermore, several science organisations focus on the development, acceptance and information exchange of 3Rs models.
Education on 3Rs models cannot be a one-time event, but should be a continuous process that educates scientists also on new developments that can be applied to replace, reduce and refine animal experimentation.

Keywords: 3Rs, continuous education, attitude

Introduction

The consideration of replacement, reduction and refinement (3Rs) of animal experiments is nowadays regarded as essential, both for improvement of research and animal welfare. "The successful implementation of the Three Rs heavily depends upon the education and training of those persons involved in research and testing" (Balls et al., 1995). Education and training should be focussed on the existing and newest advancements, but also at the development of an attitude, including the consideration of 3Rs methods, towards experimental animals.

Education is a two-party process; it involves teaching and learning, with the aim to get a change in behaviour and sometimes attitude. In general, this is a "voluntary" process, which takes place during the education at schools and universities. Learning, and change of behaviour, can also be obtained by an interaction with the environment, which may not always be on a voluntary basis.

With regard to the implementation of the 3Rs, the most effective way to educate scientists is during their study in one of the biomedical sciences. Many developments, new techniques, strategies and legislation may contribute to the 3Rs. Therefore, regular updating or continuous education on recent developments of scientists should receive more attention in order to get a better implementation of the 3Rs.

This paper examines opportunities for continuous education.

Students

Although in some countries animal experiments are still performed at primary and high schools, in most cases students are exposed to these types of experiments for the first time during their study in one of the biomedical sciences.

These experiments are generally performed under controlled conditions and supervised by skilled teachers (see also Teachers). The study should reflect the common practice with regard to experimental animals: these are not always an obvious choice but should only be used when absolutely necessary to reach the learning goals. In all other cases, alternatives should be considered and used. It is the experience that, in general, students are very conscious and critical about animal experimentation during their study and are willing to listen to opinions of different stakeholders in order to form their own opinion and attitude. In addition, they have taken initiatives to inform each other on available alternatives in education (Jukes and Chiuia, 2003).

New scientists

EU Directive 86/609/EEC (Anon, 1986) states that only competent persons, having appropriate education and training shall perform animal experiments (Article 7.1 and Article 14). To harmonise the requirements on education and training at the European level, the Council of Europe (Anon, 1994) and FELASA (Wilson et al., 1995) have proposed more detailed

provisions for a course on laboratory animal science (LAS). After completion, scientists are qualified, but not yet experienced, to design and perform animal research. The LAS course should cover several aspects related to animal research, including ethics and alternatives. The total duration of the course should take at least 80 hours.

Particularly with regard to alternatives, the following subjects should be included in the course (Wilson et al., 1995):

- Defining alternatives;
- Replacement, reduction or refinement of animal use;
- Survey of alternatives;
- Possibilities and limitations of alternatives;
- Alternatives in education and research.

In several countries, the completion of FELASA-type courses is now compulsory to qualify for directing and designing animal experiments. To further contribute to the harmonisation of the laboratory animal science courses, FELASA has established an accreditation system (Nevalainen et al., 2002).

Established scientists

After the completion of the laboratory animal science course, scientists who qualify to perform animal research normally do not automatically get updated on the progress of 3Rs methods for the rest of their career. Still, these are the persons that design the experiments. Some of them have the attitude to consider the 3Rs when designing an experiment, others may loose their original dedication to the 3Rs, while a third group may still think that animal experiments are the only way. Though it may seem difficult to directly inform these scientists, there are several ways to continuously make them aware of the importance of considering the 3Rs when designing an experiment. Continuous education can be obtained through in-house training courses (Suzuki, 2005), or by external factors that will be explored in more detail below.

Animal ethics committees

In most European countries committees that evaluate proposed animal experiments have been established, either on a compulsory, by legislation, or voluntary basis. Their names may vary from Animal Experiments Committees to Institutional Animal Care and Use Committees, but are generally referred to as animal ethics committees (AEC). Some are nation-wide, others are linked to institutions. Although their affiliation and composition may be different, they evaluate protocols for proposed animal experiments and discuss whether all 3Rs have been considered. In several countries, an experiment can only start when there is a positive advice or decision from an AEC.

Although not established to educate scientists, these committees can have a strong influence on scientists. First of all, by carefully chosen questions in the application form, the scientists are forced to critically assess their proposed experiment in all aspects and to reconsider the planned use of animals. Most committees request information on whether 3Rs models (alternatives) have been considered and why these cannot be applied. The most difficult item to be assessed by the committee is whether the potential benefit of the experiments outweighs the potential suffering of the animal.

Some of the aspects that should also be assessed by the AEC and on which the scientists should give detailed information are the level of painful procedures and methods to relieve suffering, the applied statistics to use the minimum number of animals to reach significant results, and the humane endpoints that will be applied to minimise animal suffering.

Symposia

One way to inform scientists on recent developments in the 3Rs is by organising focussed meetings. Unfortunately, organising symposia on 3Rs methods generally attract people that are already aware of the importance of the 3Rs. The persons at whom the symposia are really aimed do often not participate in these events. To inform them on the recent developments on the 3Rs, they have to be met in their own environment. Many scientific meetings are organised each year on topics that involve animal experimentation. During these events, sessions, lectures and demonstrations could be organised to raise awareness on the 3Rs.

The European Resource Centre for Alternatives in higher education (eurca (de Boo et al., 2004)) is an organisation that actively promotes alternatives in education at various national and international scientific meetings, by demonstrating animal-free teaching methods and by giving presentations on these subjects.

At the national level the responsibility for this type of information exchange could be taken by national organisations on alternatives like: NCA (Nl), REMA (Sp), SSCT (Scandinavia), ZEBET (G) and NC3RS (UK).

At the international level these activities could be organised by ECVAM, IVTIP, ESTIV and ECOPA.

Journal policies

To most scientists it is crucial to have their results published in a respectable scientific journal. Journals have set quality criteria that manuscripts have to meet for publication. Generally, these criteria cover only scientific and format aspects. Most journals that publish research on animals require from authors that a statement be made on the ethical use of animals. This requirement should be clearly stated in the author's instructions. Furthermore, journal referees should be instructed to assess whether the 3Rs have been satisfactorily considered in the submitted manuscript.

The first activities in this area were taken during the second World Congress on Alternatives and Animal Use in the Life Sciences in 1996 (van Zutphen and Festing, 2000). Out of the 47 journals scanned, 25 had no policy with regard to 3Rs (Boisvert, 1997). Guidelines were prepared and sent to editorial boards of journals that publish animal-based research (FRAME, 1999; van Zutphen and Festing, 2000).

The aspects that should at least be dealt with are:

- Statement on the followed ethical procedures;
- Justification of species and number of animals;
- Clear description of applied methods to keep animal suffering to an absolute minimum.

During a recent scan, 12 of the 25 journals do still not require a statement from the submitting authors on the ethical aspects of animal research.

Policies science organisations

Most, if not all, countries have established national science organisations. These organisations have their own research institutes and strongly influence basic research performed in their countries. In Europe, the European Science Foundation (ESF) is the umbrella organisation for the national science organisations. Through ESF policies, and the subsequent endorsement by the member organisations, there can be a direct influence on research. In 2000, the ESF adopted the position paper on the "Use of animals in research". In this document, the ESF strongly supports the principles of the "Three Rs" (Foundation, 2000; van Zutphen, 2004). As a consequence, efforts ought to be taken to replace the use of live animals by non-animal alternatives, to use the minimum number of experimental animals that is required for obtaining meaningful results and to refine procedures, so that the level of pain and suffering is minimised.

The ESF postion paper gives a strong signal to its member organisations to apply 3Rs methods. Furthermore, science organisations that endorsed the position paper can now be held accountable for the consequences of this endorsement.

Policies funding organisations

Most research projects are funded by external resources. Funding organisations preferably fund high-quality research. They could play an important role in the implementation of the 3Rs by having submitted protocols involving animal experimentation evaluated by an AEC before taking them into consideration for funding. The consideration of the 3Rs should be one of the criteria for funding.

One of the biggest funding organisations, the European Commission (EC), also funding animal experiments, does require ethical evaluation of experiments (http://europa.eu.int/comm/research/conferences/2005/recs/index_en.htm).

The EC, as funding organisation, does require "that research activities would not contravene fundamental ethical principles". However, the ethical principles of the EC only refer to humans used in research, not to animals.

When funding scientific research involving animals, the EC should at least state that the principles laid down in Directive 86/609 EEC (Anon, 1986) and the European Convention (Anon, 1991) should be respected.

With the Seventh Framework Programme to be started in the near future, a clear policy of the EC when funding research programs that likely involve animals is required. This policy should at least require an ethical review of proposed animal experiments.

National and international regulations

In the end, the most powerful way to have the 3Rs implemented is by national regulations. At the European level, both EU and European convention have established guidelines on animal experimentation.

With regard to the 3Rs the European Directive states:

- An experiment shall not be performed if another scientifically satisfactory method of obtaining the results sought, not entailing the use of an animal, is reasonably and practically available (Art. 7 {2}).
- In a choice between experiments, those which use the minimum number of animals, cause the least pain, suffering, distress or lasting harm and which are most likely to provide satisfactory results shall be selected (Art. 7 {3}).

The EU directive 86/609/EEC (Anon, 1986) is currently being revised to meet current developments. It is of utmost importance that the EU countries and those that have ratified the European Convention implement the guidelines in the national legislation.

Moreover, to be effective, every country should also have an active inspectorate in place.

Teachers

Most scientists also have responsibilities in teaching students with regard to the implementation of the 3Rs. In the end, university teachers contribute to the students' attitudes toward experimental animals. By their lecturing and choice of laboratory classes, they demonstrate the common practice with regard to experimental animals and this way contribute to the development of the students' attitude.

Since most teachers are also scientists, their awareness of the 3Rs can be raised through most of the above-mentioned activities.

To contribute to the students' attitude the teachers should:

- Make sure that the learning goals are clearly defined;
- Choose the optimal learning model, dependent on careful analysis of the learning goals;
- Always use alternatives when possible;
- Keep invasive animal experiments to an absolute minimum;
- Give an introductory lecture in case animal experiments are regarded as the only way to teach the learning goals, to discuss current ethics with regard to animal use and substantiate why for this purpose experimental animals will be used;
- Offer additional non-animal alternatives;
- Contribute to the attitude formation of students with regard to experimental animals.

Conclusion

Informing (future) scientists on the 3Rs is not only a matter of educating students in biomedical sciences. Several activities may raise the awareness on the 3Rs of scientists who are not directly involved in the field of the 3Rs. These activities should receive more attention. Experience shows that "frapper toujours" i.e. confronting scientists with the 3Rs whenever possible, will eventually result in a change of behaviour and may result in a change of attitude. The responsibility of these initiatives is not only with national and European government, but also with national and international 3Rs organisations.

In the end, both science and research animals will benefit.

References

Anon. (1986). EEC Council Directive 86/609 on the approximation of laws. Regulations and administrative provisions of the Member States regarding the protection of animals used for experimental and other scientific purposes. *Official Journal of the European Commission, L 358*, 1-29.

Anon. (1991). European Convention for the Protection Vertebrate Animals used for Experimental and other Scientific Purposes, CETS No.: 123. *Council of Europe. Retrieved 3-02-2005*, 2005, from the World Wide Web: http://conventions.coe.int/treaty/en/Treaties/Word/123.doc.

Anon. (1994). Resolution on education and training of persons working with laboratory animals. Council of Europe Multilateral Consultation. *ICLAS News, 4*, 5-6.

Balls, M., Goldberg, A. M., Fentem, J. H. et al. (1995). The three Rs: the way forward: the report and recommendations of ECVAM Workshop 11. *ATLA, 23(6)*, 838-866.

Boisvert, D. (1997). Editorial policies and animal welfare. In L. F. M. v. Zutphen and M. Balls (eds.), *Animal alternatives, welfare and ethics* (399-410). Amsterdam: Elsevier.

de Boo, M. J., Dewhurst, D., and van der Valk, J. (2004). The European resource centre for alternatives in higher education. In M. Balls, D. Firmani and A. N. Rowan (eds.), *The Three Rs at the Beginning of the 21st Century*. *ATLA 32 (Suppl. 1)*, 603-605).

Foundation, E. S. (2000). Use of animals in research. *ESF Policy Briefing, 15*, 1-6.

FRAME. (1999). Proposed guidelines for scientific journals publishing papers involving the use of laboratory animals. *Retrieved 03-10-2005*, 2005, from the World Wide Web: http://www.frame.org.uk/reductioncommittee/journalguidelines.htm.

Jukes, N., and Chiuia, M. (2003). *From guinea pig to computer mouse*. Leicester, UK: InterNICHE.

Nevalainen, T., Blom, H. J. M., Guaitani, A. et al. (2002). FELASA recommendations for the accreditation of laboratory animal science education and training - Report of the Federation of European Laboratory Animal Science Associations Working Group on Accreditation of Laboratory Animal Science Education and Training. *Lab. Anim., 36(4)*, 373-377.

Suzuki, M. (2005). Education system in Japanese research laboratories for animal experimentation in a global pharmaceutical company. *ALTEX, 22 (Special Issue)*, 31.

van Zutphen, B. F. (2004). The European Science Foundation: Three Rs policy and education. *ATLA, 32, Suppl. 1*, 533-537.

van Zutphen, B. F., and Festing, M. F. (2000). Update on policies of journals. In M. Balls, A.-M. van Zeller and M. E. Halder (eds.), *Progress in the reduction, refinement and replacement of animal experimentation* (Vol. 31B, 1679-1700). Amsterdam: Elsevier.

Wilson, M. S., Berge, E., Maess, J. et al. (1995). Felasa Recommendations on the Education and Training of Persons Working with Laboratory-Animals – Category-A and Category-C. *Lab. Anim. 29(2)*, 121-131.

Acknowledgements

Marjolein van Boxel is thanked for the critical reading of the manuscript.

Correspondence to

Dr. Jan van der Valk
Netherlands Centre Alternatives to Animal Use (NCA)
Dept. Animals, Science and Society
Fac. Veterinary Medicine
Utrecht University
Yalelaan 2, P.O. Box 80.166
3508 TD Utrecht
The Netherlands
e-mail: b.f.vandervalk@vet.uu.nl

The Use of Animals in Research, Testing and Teaching in New Zealand – A Legal Perspective

Neil Wells
School of Natural Sciences, Unitec, New Zealand

Summary
Part 6 of the Animal Welfare Act 1999 (New Zealand) provides that a person (which by definition includes an institution) may only use an animal for the purposes of research, testing or teaching if the person is a code holder or is authorised by a code holder, the research, testing and teaching is approved by an institutional animal ethics committee and carried out in accordance with any conditions imposed by the AEC.
Legal provisions such as the Animal Welfare Act are not intended to be, nor should they be used, as an insurance policy. The concept of AECs approving animal use protocols "just in case" without careful inquiry is an abuse of legal process and may be unnecessarily time consuming and costly for the applicant, the code holder and the animal ethics committee. It is critical that all the legal requirements of the Act are met.
Part 6 of the Animal Welfare Act 1999 embodies the legal provisions of the 1984 amendment to the Animals Protection Act 1960 and practices that developed until 1998. Where Part 6 of the Act does not apply to a particular animal use Part 1 will apply. This paper focuses on determining which part of the Act applies to specific animal use in research, testing and teaching and provides a guide to assist AECs in determining when approval under Part 6 of the Act is required.

Keywords: legislation, manipulation, teaching, animal use

Introduction

On 1 January 2000 the Animal Welfare Act 1999 (New Zealand) commenced. Part 6 of the Act provides the framework for regulating the use of animals in research, testing and teaching. This Part of the Act embodies the legal provisions of the 1984 amendment to the Animals Protection Act 1960 together with practices that developed until 1998 when the Animal Welfare Bill (No. 2) was drafted. Where Part 6 of the Act does not apply to a particular animal use in research, testing and teaching, Part 1 applies.

This paper is based in the principles of legal interpretation, it is not a scientific paper, nor is it based on an ethical analysis. In applying the tenets of legal interpretation to the use of animals in research, testing and teaching, the provisions of the Animal Welfare Act 1999 must be considered as a whole as well as the purposes of Part 6.

Part 1 establishes a duty of care for animals. The owner and every person in charge of an animal must provide for its physical, health and behavioural needs.[1] Person in charge, "in relation to an animal, includes a person who has that animal in that person's possession or custody, or under that person's care, control or supervision."[2]

Part 6 of the Act deals exclusively with animals used in research, testing and teaching. While this Part has sometimes been regarded as a "stand alone" Part, it must still be read in the context of the whole Act. Before an animal can be said to be "used" in research, testing and teaching there must be a "manipulation". The definition of research, testing and teaching is implicit in that the term means:

(a) Any work (being investigative work or experimental work or diagnostic work or toxicity testing work or potency testing work) that involves the manipulation of any animal; or
(b) Any work that
  (i) Is carried out for the purpose of producing antisera or other biological products; and
  (ii) Involves the manipulation of any animal; or
(c) Any teaching that involves the manipulation of any animal.[3]

In each instance the work or teaching is qualified as involving a manipulation, which means interfering with the normal physiological, behavioural, or anatomical integrity of the animal by deliberately–

(a) Subjecting it to a procedure which is unusual or abnormal when compared with that to which animals of that type would be subjected under normal management or practice and which involves–
  (i) Exposing the animal to any parasite, micro-organism, drug, chemical, biological product, radiation, electrical stimulation, or environmental condition; or
  (ii) Enforced activity, restraint, nutrition, or surgical intervention; or
(b) Depriving the animal of usual care.[4]

Legislation is not intended to be used as an insurance policy on a "just in case" basis. The practice of treating all animal use in research, testing and teaching as a manipulation without careful inquiry is not only an abuse of process, i.e. using regular legal process for a purpose not contemplated in law, it involves

[1] Animal Welfare Act 1999, section 10
[2] section 2
[3] section 5
[4] section 3

unnecessary time and cost to the researcher, tester or teacher, to the institutional animal ethics committee, to the code holder, to the Ministry of Agriculture and Forestry in collating statistical records, and time and cost of reviews. The Animal Welfare Act 1999 provides a legal framework for the use of animals in research, testing and teaching.

Legal interpretation is a discipline that seeks to interpret the words and terms used by Parliament and examines what was intended by Parliament. While it is not an exact discipline it is governed by legal precedent and statute. Shortly before the Animal Welfare Act 1999 was passed the Interpretation Act 1999 commenced. The method to be applied to legal interpretation is unequivocal:

Ascertaining meaning of legislation

(1) The meaning of an enactment must be ascertained from its text and in the light of its purpose.
(2) The matters that may be considered in ascertaining the meaning of an enactment include the indications provided in the enactment.
(3) Examples of those indications are preambles, the analysis, a table of contents, headings to Parts and sections, marginal notes, diagrams, graphics, examples and explanatory material, and the organisation and format of the enactment.[5]

Part 6 states its purposes in section 80 where it is set out that its principal purpose is to ensure that, *inter alia*, the use of animals in research, testing and teaching is confined to cases where there is good reason to believe the findings and benefits will enhance understanding of human beings, animals, or ecosystems or the maintenance and protection of human or animal health or welfare.[6]

A further purpose includes ensuring that animals used in research, testing and teaching have their physical, health and behavioural needs met. More specifically, a purpose of this Part can be found in the expression of the Three Rs.

In 1959 Russell and Birch first published *The Principles of Humane Experimental Technique*[7]. The authors considered "the ways in which inhumanity can be and is being diminished or removed" in experimental technique by–

Replacement – the substitution for conscious living higher animals of insentient material

Reduction – reduction in the numbers of animals being used to obtain information of a given amount and precision

Refinement – any decrease in the incidence or severity of inhumane procedures applied to those animals which still have to be used.[8]

The Act echoes these principles in that a further purpose of Part 6 is –

(b) To promote efforts –

(i) To reduce the number of animals used in research, testing, and teaching to the minimum necessary:

(ii) To refine techniques used in any research, testing, and teaching so that the harm caused to the animals is minimised and the benefits are maximised:

(iii) To replace animals as subjects for research, and testing by substituting, where appropriate, non-sentient or non-living alternatives:

(iv) To replace the use of animals in teaching by substituting for animals, where appropriate, non-sentient or non-living alternatives or by imparting the information in another way.[9,10]

Part 6 of the Act uses the words "use" and "manipulation" interchangeably. The heading of Part 6 is the "Use of Animals in Research, Testing and Teaching". Using the rule of interpretation of the Interpretation Act 1999, the heading clearly means that in Part 6 the term "use" means "manipulation" as "manipulation" is implicit in the definition of "research, testing and teaching".

The term "use" is not defined by the Act, but using the *noscitura a sociis* rule, "a word is known by its context" or "a word may be known by the company it keeps", the term "use" must take its meaning from its context. In Part 6 it can only mean "manipulation" rather than the wider common meaning of "the act of using a thing for a purpose … utilisation or employment for or with some aim or purpose".[11] A group of animals used to explain anatomy or physiology without physical intrusion into their bodies is a use of animals for the purpose of teaching – but it is not necessarily a manipulation. Similarly other benign uses of animals in teaching where there may be some physical handling may not be a manipulation, e.g. teaching veterinary nurses bandaging techniques uses animals in teaching but is hardly a manipulation.

While the term "project"[12] does not refer to research, testing and teaching, nor to animals, the term still needs to be read in the context that the whole of Part 6 is concerned with research, testing and teaching, so it is patently obvious that it does relate to animals. It would be absurd to say that because the definition of a "project" means *inter alia* "any experiment or series of related experiments, forming a discrete piece of research" means that a physics or engineering experiment requires animal ethics committee approval when clearly the term does not require that.

Despite the practice that all use of live animals in a research, testing or teaching project must gain approval from an animal ethics committee that is not the case in all circumstances of animal use, particularly when it is a benign use for teaching purposes. While by far the majority of animal use will require institutional animal ethics committee approval it is only when a number of interconnected elements are present that the use of an animal becomes a manipulation and thus comes within the scope of Part 6 of the Act. These elements can be summarised as:

---

5 Interpretation Act 1999 section 5

6 Animals Protection Act 1960 section 80 (1)

7 Russell, W. M. S. and Burch, R. L. (1992). *The Principles of Humane Experimental Technique*, special edition. Universities Federation of for Animal Welfare, Potters Barr.

8 Russell & Burch, 64

9 section 80 (2) (b)

10 section 5 (3)

11 *The Shorter Oxford Dictionary On Historical Principles* (1973). Oxford: Clarendon Press, 2441.

12 section 2

- The species of animal and the living state of that animal
- The purpose for which the animal is used
- Whether or not the animal is manipulated.

Before considering the first step there is a presumption that the applicant has considered non-animal alternatives to the proposal in the spirit of the Three Rs and that this prior consideration can be demonstrated to the institutional animal ethics committee.

The species of animal

Only animals as defined by the Act are subject to Part 6. The Animal Welfare Act 1999 provides that –

"Animal"–

(a) Means any live member of the animal kingdom that is –
  (i) A mammal; or
  (ii) A bird; or
  (iii) A reptile; or
  (iv) An amphibian; or
  (v) A fish (bony or cartilaginous); or
  (vi) Any octopus, squid, crab, lobster, or crayfish (including freshwater crayfish); or
  (vii) Any other member of the animal kingdom which is declared from time to time by the Governor-General, by Order in Council, to be an animal for the purposes of this Act; and

(b) Includes any mammalian foetus, or any avian or reptilian pre-hatched young, that is in the last half of its period of gestation or development; and

(c) Includes any marsupial pouch young; but

(d) Does not include –
  (i) A human being; or
  (ii) Except as provided in paragraph (b) or paragraph (c) of this definition, any animal in the pre-natal, pre-hatched, larval, or other such developmental stage:[13]

The Act applies to all vertebrates and some invertebrates; it applies to the foetus of any mammal or the pre-hatched young of any bird or reptile, provided it is in the second half of development; humans are excluded.

The animal needs to be living, in other words animal tissue or dead animals do not come within the scope of Part 6 of the Act but the method of obtaining the dead animal may still come within the scope of Part 1.

The purpose for which the animal is used

The next step is to determine the purpose for which the animal is being used. Is it to be used for research, testing or teaching? Research is defined as:

(a) Any work (being investigative work or experimental work or diagnostic work or toxicity testing work or potency testing work) that involves the manipulation of any animal.[14]

Testing is defined as:

(b) Any work that –
  (i) Is carried out for the purpose of producing antisera or other biological products; and
  (ii) Involves the manipulation of any animal.[15]

Teaching is defined as:

(c) Any teaching that involves the manipulation of any animal.[16]

This last definition still does not define "teaching", which can best be described as "to show by way of information or instruction; to impart or convey knowledge of; to give instruction or lesson in (a subject)."[17] There should be some structure to the imparting of information before it can be said to be teaching, such as in a class or tutorial group. It can hardly be said that a farmer showing a farm hand the correct method of milking a cow is "teaching" in the context of the Act.

Whether or not the animal is manipulated

There are two conditions that need to be considered before it can be said that an animal is "manipulated".

First, there must be some interference with the normal physiological, behavioural or anatomical integrity of the animal by deliberately subjecting it to a procedure that is unusual or abnormal when compared to that which animals of that type would be subjected under normal management or practice and which involves exposing the animal to any –

- Parasite
- Micro-organism
- Drug
- Chemical
- Biological product
- Radiation
- Electrical stimulation
- Environmental condition
- Or enforced activity, restraint, nutrition or surgical intervention.[18]

Alternatively, there must be some interference with the normal physiological, behavioural or anatomical integrity of the animal by deliberately depriving the animal of usual care.

"Interference" is not defined by the Act but "to interfere" can be understood to mean "to obstruct a process, to interpose so as to affect some action, to meddle with."[19]

One of the Three Rs reflected in section 80 (2) (b) is–

(ii) To refine techniques used in any research, testing, and teaching so that the harm caused to the animals is minimised and the benefits are maximised:

Some assistance can be gained from clause (ii) where it refers to potential "harm caused to animals". One test as to whether a particular interference can be said to be interference in the physical, behavioural or anatomical integrity of an animal is to determine whether or not the animal will be harmed. If harm is a likelihood then clearly there is a manipulation. The converse implies that if the use will cause no harm to the animal, or there is no real likelihood that harm will be caused, then it may not be a manipulation at all. But each project needs to be assessed by

[13] section 2
[14] section 5(1)(a)
[15] section 5(1)(b)
[16] section 5(1)(c)
[17] *Shorter Oxford Dictionary,* 2550
[18] section 3 (1) (a)
[19] *Shorter Oxford Dictionary,* 1094
[20] section 4

the institutional animal ethics committee on a case-by-case basis before it can be said that there is no manipulation. Where there is no manipulation Part 6 will not apply.

Where a particular use of an animal has been determined not to be a manipulation under Part 6, the owner of the animal, usually the institution, and the person in charge of the animal, each have a legal responsibility under Part 1 to ensure that the physical, health and behavioural needs of the animal are met.[21] In particular the provision of –

(a) Proper and sufficient food and water;
(b) Adequate shelter
(c) Opportunity to display normal patterns of behaviour:
(d) Physical handling in a manner which minimises the likelihood of unreasonable or unnecessary pain and distress:
(e) Protection from, and rapid diagnosis of, any significant injury or disease, –

being a need which, in each case, is appropriate to the species, environment, and circumstances of the animal.[22]

As a further aid the applicant must be able to express the intent of the project for it is the deliberate subjection of the animal to physiological, behavioural, or anatomical intervention that determines whether or not a use is a manipulation. For example, in teaching a veterinary student how to intubate an animal there is a clear intention to submit the animal to a teaching project that will interfere with the physical integrity of the animal by surgical intervention and therefore it is a manipulation. On the other hand the same student may be taught the correct way to pick up and hold a cat but in the process drops it. There is no intention in the project to drop the animal so it is not a manipulation, which must involve "deliberately subjecting it to a procedure."[23] Nonetheless, there will be some other provision of the Act, such as the requirement to ensure physical handling in a manner which minimises the likelihood of unreasonable or unnecessary pain or distress,[24] which would apply to the negligent handling of an animal. Similarly, if a group of laboratory animals was deprived of nutrients as part of an experimental project, clearly that is a manipulation. But if laboratory animals were deprived of food due to negligence, it is not a manipulation but a failure to provide for the physical, health and behavioural needs of the animals.

The test then as to whether or not an animal is being manipulated is in the analysis of each element of manipulation. Unless all elements are present it is not a manipulation. In summary, before it can be said that there is a manipulation there must first be an intention to use an animal for research, testing or teaching, and secondly there must be an interference with the normal physiological, behavioural or anatomical integrity of the animal by deliberately subjecting it to a procedure that is unusual or abnormal, or by deliberately depriving it of its usual care. If any one element is missing the animal use is not a manipulation.

Exceptions

An animal is not deemed to be used for the purposes of research, testing or teaching, and therefore no code of ethical conduct is required, where an animal is in the immediate care of a veterinarian and the veterinarian believes on reasonable grounds that the manipulation will not cause the animal unreasonable or unnecessary pain or distress, or lasting harm; and the manipulation is–

(i) For clinical purposes in order to diagnose any disease in the animal or any associated animal; or
(ii) For clinical purposes in order to assess the effectiveness of a proposed treatment regime for the animal or any associated animal; or
(iii) For the purposes of assessing the characteristics of the animal with a view to maximising the productivity of the animal or any associated animal.[25]

There are some specific exclusions from the definition of "manipulation":

(a) Any therapy or prophylaxis necessary or desirable for the welfare of an animal; or
(b) The killing of an animal by the owner or person in charge as the end point of research, testing, or teaching if the animal is killed in such a manner that the animal does not suffer unreasonable or unnecessary pain or distress; or
(c) The killing of an animal in order to undertake research, testing, or teaching on the dead animal or on pre-natal or developmental tissue of the animal if the animal is killed in such a manner that the animal does not suffer unreasonable or unnecessary pain or distress; or
(d) The hunting or killing of any animal in a wild state by a method that is not an experimental method.[27]

Where an animal is killed as the end point of research, testing and teaching, the act of killing is excluded from the provisions of Part 6 provided the animal is killed in a manner that does not inflict unreasonable or unnecessary pain or distress.[28] Any prior treatment or handling of the animal so killed that constitutes a manipulation continues to be covered by Part 6 and therefore does require a code of ethical conduct and the project does not need to be approved by an institutional animal ethics committee.

Further, an animal is not deemed to be used for the purposes of research, testing or teaching where the manipulation involves routine breeding, marking, capturing, translocation trapping, weighing or taking measurements of animals or work carried out with the principal objective of certain functions under the Conservation Act 1987 or the Fisheries Act 1996.[29]

Recording statistics

Statistics in the Annual Report of the National Animal Ethics Advisory Committee (NAEAC) for 2003 record that in that year 320,911 animals were manipulated in research, testing and teaching, and a range of tables that follow uses the term "use" or

[21] section 10
[22] section 4
[23] section 3 (1)
[24] section 4
[25] section 5 (2) (b)
[26] section 5 (3)
[27] section 3 (2)
[28] section 3 (2) (b)
[29] section 5 (3)

"usage". It is noted that the published statistics record "animal usage" rather than "animal manipulations". It is possible that these statistics have been overstated by the reporting of "animal usage" that was not by definition "manipulation". Regulation 5 of the Animal Welfare (Records and Statistics) Regulations 1999 requires the reporting of information recorded under regulation 4 by requiring every code holder to record information relevant to animals manipulated.

While section 183 enables regulations requiring every code holder to keep information in relation to inter alia the numbers and species of animals used , the word "used" must be interpreted by the noscitura a sociis rule, and also by the definition of "code holder"[32] as being a person carrying out research, testing and teaching, which by definition requires a manipulation to be present. In this context the term "used" is incapable of meaning "any use" whether or not there is a manipulation. There is no requirement to record or report information relevant to animals used where there was no manipulation. Clearly the statistics are intended to refer to manipulations.

The 2003 statistics include 78,520 (24.4%) that were reported as "no suffering" and 195,451 (60.1%) were reported as "little suffering".[33] – some, particularly in the "no suffering" class, may well have fallen short of the threshold of being manipulated. If that were the case the statistics may be over-reported. Unless each institutional animal ethics committee is carefully inquiring into the precise "use" of animals in each application, "uses" that do not meet the threshold of "manipulation" may be inadvertently included in the reported statistics.

How the reported use of animals is classified in terms of its impact is dealt with in depth by a contemporary paper by Williams, Mellor and Marbrook presented in the 5th Congress session *Ethical Review – Good Practice* and Outputs.[34]

Conclusion

Institutional animal ethics committees are charged with approving or not approving projects that use animals in research, testing and teaching.[35] Institutional animal ethics committees should also, as part of that process, consider carefully whether an application under consideration does in fact propose to manipulate any animal. By applying these guidelines it may well be that many uses of animals in research, testing or teaching will not be classed as manipulations at all and, thus, institutional animal ethics committees will be saved the task of approving and monitoring projects and recording statistics relating to projects for uses that the Animal Welfare Act 1999 does not intend to come within the scope of Part 6.

However, it is not intended that this paper be used as a convenient legal loophole for researchers, testers and teachers. For the protection of the integrity of both the applicant and the institute it is strongly recommended that where an animal is to be used in research, testing or teaching, and where no manipulation to the animal is apparent, the institutional animal ethics committee should make careful inquiry into the nature of the application and its intent and record its decision as to whether or not the proposed use is a manipulation. Should it be determined that the application does not involve a manipulation of an animal, the owner of the animal (the institution) and every person in charge of the animal still have a legal duty to ensure that the physical, health and behavioural needs of the animals are protected.[36]

Acknowledgements

I acknowledge the helpful and sometimes challenging comments of David Mellor (Massey University) and David Bayvel, Linda Carsons, and Kate Littin (Biosecurity New Zealand) – this paper is the better for their input. However, responsibility for the opinions expressed herein rest with me.

Correspondence to

Neil Wells BA LLB (Hons), Barrister
Associate Head of School,
School of Natural Sciences
Unitec, New Zealand
e-mail: nwells@unitec.ac.nz

[30] Annual Report of the National Animal Ethics Advisory Committee, 2003

[31] section 183 (1) (c)

[32] section 183 (2)

[33] Annual Report of the National Animal Ethics Advisory Committee, 2003

[34] Williams, V., Mellor, D. and Marbrook J. (2005). *Revision of a Scale for Assessing the Severity of Live Animal Manipulations*

[35] section 99

[36] section 10

Workshop 1.4
Multi-media exhibition of alternatives in education

The NORINA & TextBase Website: New Design and Possibilities

Karina Smith[1], Adrian Smith[1], Baard Johannessen[2] and Arne Lie-Johannessen[2]
[1] Norwegian School of Veterinary Science, Laboratory Animal Unit, Oslo, Norway
[2] Nettforsk, Arendal, Norway

Summary
The Norwegian Reference Centre for Laboratory Animal Science & Alternatives has maintained a website since 1996, featuring the NORINA database (http://oslovet.veths.no). NORINA contains information on nearly 4,000 products that can be used as animal alternatives or supplements in teaching and training. The website includes TextBase, a database with information on over 1,000 textbooks of relevance to the three R's. The website has been totally rebuilt in 2005. The databases are now part of one relational database, including information on guidelines for animal research, current legislation, course material, other databases and a Virtual Tour of the Centre.

Keywords: NORINA, TextBase, website, alternatives, teaching

Introduction

The interest in alternatives to the use of animals in research and education created a need for specialist information sources addressing the Three Rs of replacement, reduction, and refinement (Russell and Burch, 1959). In many parts of the world there are mandatory requirements for the use of alternatives where practically possible, as exemplified by the EU Directive 86/609 (EEC, 1986) and the Council of Europe Convention ETS 123 (Council of Europe, 1986).

Description of the website

The website of the Laboratory Animal Unit at the Norwegian School of Veterinary Science was created to meet these demands, by presenting free of charge on the Internet a database of audiovisual aids and other materials that can be used as pure alternatives, or supplements, to animals or animal materials in teaching and training, at all levels from junior school to University (http://oslovet.veths.no). This database, known as NORINA (A Norwegian Inventory of Alternatives) was subsequently followed up by a satellite database (TextBase) containing information on textbooks and other written resources that may also be used in teaching and training, but that also have a role to play in the general refinement of laboratory animal science at any facility. NORINA currently contains information on nearly 4,000 products, while TextBase covers over 1,000 textbooks and other publications.

The development of the website coincided with a workshop organised by ECVAM in 1996 (Janusch et al., 1997) which concluded that information on existing databases adressing the three Rs should be systematised and made readily available. These recommendations were further developed within the field of education by another ECVAM workshop, in 1998 (van der Valk, 1999). Therefore, the Oslovet site was expanded to include information on other databases around the world, together with links to websites and textual resources within this field (fig. 1).

The Oslovet website has been totally rebuilt in 2005. All the information in NORINA and TextBase, together with the other materials on the website (webpages, documents and images) has been incorporated into one large relational database using state-of-the-art asp.net technology. This allows users, among other things, to search the whole website simultaneously, and then to access

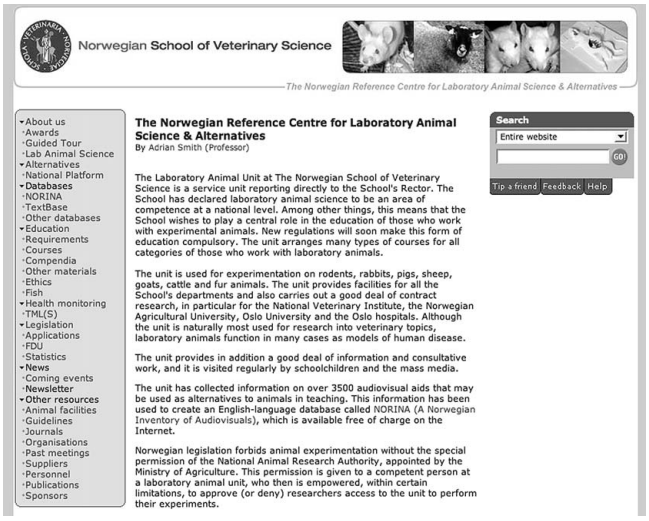


Fig. 1: An overall view of the website of the Norwegian Reference Centre for Laboratory Animal Science and Alternatives.

search results from NORINA, TextBase or the other material, depending upon their needs. It also enables the database compilers, and other authorised editors, to update any section of the website instantaneously via the Internet from any location in the world.

A number of new features have been added, to take advantage of the new technology. These include:

- direct links to suppliers, publishers, Internet bookstores and Google, so that searches can be repeated there with no further typing required;
- lists of new products, reviewed items, materials held at the Centre, items which the Centre considers to be "key products", and materials that are available free of charge or on loan from international loans programs;
- "Tip a Friend", "Help", "Feedback", "News" and "Newsletter" functions;
- a comprehensive statistics package giving a detailed breakdown of all activity at the site.

In addition to the NORINA and TextBase, the website contains a comprehensive collection of resources and information material within laboratory animal science in general, and the three R's in particular. These include:

- Guidelines on the care and use of fish in research, including material from an international consensus meeting held in Oslo in May 2005 (Smith et al., 2005);
- Links to other Guidelines on the care and use of animals in research and teaching (e.g. Smith and Smith, 2004; Smith and Allen, 2005);
- An overview of databases worldwide that address the three Rs;
- Teaching material for use in laboratory animal science;
- A Guided Tour of the Centre in three languages;
- General information and links for students conducting projects on animal experimentation.

The authors hope that this website will prove to be a valuable resource for further implementation of the concepts of Russell and Burch, and we welcome suggestions for improvement from users.

Acknowledgements

The financial support given by many organisations to the work of maintaining the NORINA and TextBase databases is gratefully acknowledged (http://oslovet.veths.no/sponsors.html). We wish also to thank Tim Morris, Paul Flecknell and Peter Nowlan for their valuable comments during the development of this website.

References

EEC (1986). Council Directive 86/609/EEC of 24 November 1986 on the approximation of laws, regulations and administrative provisions of the Member States regarding the protection of animals used for experimental and other scientific purposes. *Official Journal of the European Communities L358*, 1-29. http://europa.eu.int/smartapi/cgi/sga_doc?smartapi!celexapi!prod!CELEXnumdoc&lg=en&numdoc=31986L0609&model=guichett Accessed 23.09.2005.

Council of Europe (1986). European Convention for the Protection of Vertebrate Animals Used for Experimental and Other Scientific Purposes, ETS No. 123. Strasbourg, France. http://conventions.coe.int/Treaty/en/Treaties/Html/123.htm Accessed 23.09.2005.

Janusch, A., van der Kamp, M., Bottrill, K. et al. (1997). Current status and future developments of databases on alternative methods. The Report and Recommendations of ECVAM Workshop 25. *ATLA 25*, 411-422.

Russell, W. M. S. and Burch, R. L. (1959). *The Principles of Humane Experimental Technique*, 238pp. London: Methuen.

Smith, A. J. and Allen, T. (2005). The use of databases, information centres and guidelines when planning research that may involve animals. *Animal Welfare 14*, 347-359.

Smith, A. J. and Smith, K. (2004). Guidelines for humane education: Alternatives to the use of animals in teaching and training. Proceedings of the Fourth World Congress on Alternatives and Animal Use in the Life Sciences, New Orleans, 11-15 August 2002. *ATLA 32, Supplement 1A*, 29-39. http://www.worldcongress.net/2002/proceedings/SP4%20Smith.pdf Accessed 23.09.2005.

Smith, A. J., Johansen, R. and Knudsen, G. (eds.) (2005). Harmonisation of the Care and Use of Fish in Research, a consensus meeting held at Gardermoen, Norway, 23rd-26th May 2005. http://oslovet.veths.no/fish Accessed 23.09.2005.

Van der Valk, J., Dewhurst, D., Hughes, I. et al. (1999). Alternatives to the Use of Animals in Higher Education. The Report and Recommendations of ECVAM Workshop 33. *ATLA 27*, 39-52.

Correspondence to

Karina Smith
Norwegian School of Veterinary Science
Laboratory Animal Unit
P.O. Box 8146 Dep.
0033 Oslo, Norway
e-mail: karina.smith@veths.no

Theme 2
Laboratory animal welfare and refinement

Workshop 2.1
Environmental enrichment and housing standards

Refinement Alternative for Animal Housing – Enrichment

Timo Nevalainen[1,2], Niina Kemppinen[3] and Anna Meller[3]
[1]National Laboratory Animal Center, University of Kuopio and [2]Department of Basic Veterinary Sciences,
[3]Viikki Laboratory Animal Center, University of Helsinki, Finland

Summary
Enrichment is a refinement as it should promote animal wellbeing. Any refinement needs to be beneficial to animals and should not interfere with the study. While refinement aims can be connected to research data, reduction alternative suffers from lack of research. Instead of trying to assess the impact of enrichment on variance of every determination, it could be more productive to look at effects on variance of welfare indicators; low variance there is likely to reflect as low variance in other determinations. And at the same time aim at most uniform welfare of the animals in the study.

Keywords: environmental enrichment, refinement alternative, reduction alternative

Background

The use and welfare of laboratory animals in research are major issues in modern society. Both public concern on welfare of animals and continuously increasing demands on quality of both animals and of biomedical research make the issues urgent and complex. As a result the use of laboratory animals is experiencing the turmoil of rapidly changing regulations both at European and national levels.

Environmental enrichment is a term increasingly used in connection with laboratory animal care and housing. Baumans (1999) defines enrichment as any modification in the environment of the captive animal that seeks to enhance its physical and psychological well-being by providing stimuli meeting the animals' species-specific needs. This is a performance based definition, as are the definitions for both refinement and reduction alternatives.

Unfortunately the performance of an enrichment item or program has not always been shown before it has been taken into use. Indeed, various items may be placed into the animal enclosure and called enrichment without verification of their efficacy. Some elements in routine care, like use of bedding and group housing can be considered enrichment, and these have been in use before the term was even introduced.

Revised Appendix A of the Council of Europe Convention calls for environmental enrichment and group housing for all gregarious species unless there are scientific or veterinary reasons not to do so (Council of Europe, 2002). Enrichment is considered as refinement as it should promote animal wellbeing. Interference with an experimental outcome could be a scientific reason, and fighting between incompatible animals a veterinary reason for not implementing enrichment.

A closer look at the European regulations (The Convention and The Directive, 1986) reveals that these are tailored to primarily ensure animal welfare. Yet, these must be regarded as minimum standards, as spelled out in the revised documents of Appendix A of the Council of Europe (2002). Indeed, they address only in passing and superficially natural concerns of the scientific community, i.e. whether certain regulations may interfere with the studies.

Optimal enrichment

It would be naïve to believe that all enrichment programs have only positive effects on animals, and never interfere with research. Enrichment can be evaluated in terms of refinement and reduction. Refinement is essentially welfare status of the animal, and enrichment should improve the status by definition. reduction

potential of enrichment should be assessed at the same time, and we should prefer those, which decrease the number of animals needed or at least cause no change.

Any refinement to improve animal welfare requires scientific validation to ensure it is truly beneficial to animals (efficacy) and does not detract from the scientific integrity of the study (safety). When enrichment fails to fulfill one of these criteria, animal welfare may be challenged or the study should not have been done at all. The effect of enrichment on safety may simply be a change in the mean of the determination of interest, and this may not matter as it should affect all groups, but changes in variance will lead to more animals being used, itself an ethical issue. The opposite could also happen and results could be improved leading to fewer animals being used.

While refinement aims can mostly be connected to research data, reduction suffers from lack of research to base regulations on. It is obvious that changes in variance may be strain-, facility- and enrichment-specific, which makes overall guidelines difficult. Indeed, instead of trying to assess impact of enrichment on every determination, it could be more productive to look at effects on variance of welfare indicators with the understanding that low variance there is likely to show as low variance in other determinations. And at the same time aim at most uniform welfare of the animals in the study.

Discussion of enrichment in connection to GLP-studies emphasizes lowest possible levels of pesticides and heavy metals in enrichment items because they can be toxic to the animals. This is just part of the problem – safety of enrichment has to be understood wider; it is also compromised when any compounds that could change the results are brought into the enclosure. Volatile compounds, especially pinenes, which cause induction of liver microsomal enzymes, are a challenge to safety as well. Recent study showed that there still are pinenes present in beddings and that the same is true for enrichment items (Meller et al., 2004)

Pair-housing of rabbits and variance

For social animals, such as rabbits, another animal in the same enclosure is the most challenging enrichment factor. While objects intended for enrichment are static and may have only novelty value, a partner poses continuous and unpredictable challenges to which the individual must react (Stauffacher et al., 2001).

Rabbit as a species is difficult to house in groups, and group housed rabbits show preference for group sizes of one to three animals (Whary et al., 1993), and when housed in pairs, proximity of each other (Huls et al., 1991). Group size of two rabbits – pair-housing – is indeed the simplest form of social enrichment and possible in reasonably sized cages. In order to enhance social enrichment in rabbits, the European regulations will make it possible to have two rabbits in a cage space considered adequate for one (Council of Europe, 2002).

A study designed to simultaneously assess both compatibility and reduction outcome of rabbit pair-housing was done following the revised Council of Europe Appendix A (2002). Variances in growth and serum alkaline phosphatase were significantly smaller in pair-housed as compared to single housed rabbits. If growth is the resulting parameter, then with change from single housing to pair-housing in rabbits one could halve the number of animals needed, perhaps even more. The large decrease in variance can be due to continuous interaction between the pair keeping them busy in social activities at the expense of eating and resting. The point estimate for alkaline phosphatase shows that reduction of at least half of the female rabbit numbers can be considered as anticipated effect of pair-housing (Nevalainen et al., submitted).

Work for food and enrichment

Enrichment could also be used as "a cure" for excessive variation when the source can be identified. *Ad libitum* feeding is considered a poorly controlled factor, likely to result in large result variance and hence require unnecessarily many animals in order to obtain results of the specified significance and power (Keenan et al., 1999). Virtually all rodents are on *ad libitum* feeding. Refinement of the feeding practice of laboratory rodents is likely to improve scientific quality, animal health and welfare. However, in order to implement restricted feeding regimes in rodents, research is needed to achieve systems avoiding the negative effects and ensuring that all group-housed rodents have access to their share of the food provided.

An approach to solve the problems of *ad libitum* feeding is work for a food and enrichment hypothesis. Food is an indispensable item, without which animals cannot survive and endure. It is suggested that rodents will work only for the amount of food they necessarily need, but not beyond that, provided the intensity of the work is set correctly. A simple system has been invented, and a preliminary pilot study shows that it indeed decreases weigh gain of laboratory rats (Nevalainen et al., unpublished data). Full scale studies are needed to ultimately verify efficacy, and to set the dose of work. Safety evaluation is also a necessity, e.g. to assess whether animals are too lean, whether their time budget for working is too hard and what happens to variance of results.

Furthermore, animal well-being is important when assessing any novel feeding regimen. We hypothesize that determination of faecal corticosterone and IgA combined with telemetric cardiovascular parameters and behavior assessment the same animals will give a comprehensive picture of any change and its impact and so help to optimize animal welfare.

COST Action B24

COST Action B24 "Laboratory Animal Science and Welfare" is a new scientists' network focusing on both efficacy and safety of animal housing, including environmental enrichment, and designed to look for answers to the concerns of both the public and the scientists. The approach chosen is to increase knowledge necessary for both ethically sustainable and scientifically valid use of laboratory animals in research. The Action believes that these two aims are not only possible, but indeed an absolute necessity. The Action has planned for five working groups, and enrichment is included into the remit of Working Group 1, Housing of animals and scientific integrity.

The Action serves as an interaction podium and idea generator for scientists and civil servants and paves the way for European research consortia. Furthermore, it aims at the production of

research results and collection of technical data based on scientific studies, and ultimately seeks tools needed for real life implementation. Delivery of the processed data is done through harmonizing of training of persons working with animals and as guidelines and recommendations, which should go beyond regulatory minimum standards. This aim can be realized only if a strong and active network is present. A special aim of the Action, in order to further implementation, is a compilation of relevant guidelines and recommendations produced prior to the Action and by the Action into a format, which would allow all interested groups within the field to become familiar with the recommendations in a convenient, but efficient way.

The two Rs initiative

The Replacement alternative is simply not always possible, sometimes not even desirable, and hence there should be more research on the remaining Two Rs to help those animals still being used and the scientists. Replacement alternative has been studied far more than the other two. The COST Action B24 believes that all 3Rs should have equal weight in funding, and calls for full implementation of both the European Convention and in the Directive. To further this goal, the Action has prepared the Two Rs Initiative and submitted it to be included in the European Union's 7th Framework program.

Studies on alternative methods for applications in fundamental biomedical research receives too low funding, even though there is high potential for the 3Rs, and especially for reduction and refinement. Every scientist using animals can and should actively seek implementation of one or more of the alternatives.

The document "Science and technology, the key to Europe's future – Guidelines for future European Union policy to support research" (2004) states that the Commission has made strengthening European research a major objective…is proposing to increase the European Union's research budget….the budget should be doubled. If this is to happen, it will inevitably mean more laboratory animals used in fundamental research, and acute and urgent need for funding of studies of how best to apply the Two Rs methods.

Investing in research of the Two Rs alternatives at the same pace with funding of basic research enables Europe to maintain or even increase the lead it has. All research carried out should simply be ethically sustainable and scientifically valid. Yet, this is not simple nor straightforward, and it can only be achieved through tailor made, considerable funding granted on a competitive basis.

Any considerable funding to research on the Two Rs is likely to generate new knowledge enabling better welfare for fewer animals in research and consequently ease the concerns of the society. It can also be foreseen that this very same research, if carefully planned and executed, is crucial in avoiding practices and procedures compromising the scientific validity of research.

Standardisation of enrichment

The FELASA-working group on "Standardisation of Enrichment" was established with the goal to provide guidance of how to standardise enrichment in laboratory animal enclosures such that essential species-specific needs and individual needs of gender and life stage are fulfilled to guarantee animal welfare and to minimise interference with experimental results. It is expected that once the document will be published, the recommendations given will further the implementation of enrichment so that animal welfare and good science are promoted.

References

Baumans, V. (1999). Environmental enrichment: practical applications. In: L. F. M. van Zutphen and M. Balls (eds.), *Animal Alternatives, Welfare and Ethics* (187-191). The Netherlands: Elsevier BV.

Commission of the European Communities (2004). *Science and technology, the key to Europe's future – Guidelines for future European Union policy to support research.* Brussels 2004.

Council Directive 86/609/EEC of 24 November 1986 on the approximation of laws, regulations and administrative provisions of the Member States regarding the protection of Animals used for Experimental and Scientific purposes.

Council of Europe (2002). *Draft proposal for the revised Appendix A – guidelines for accommodation and care of animals*. Strasbourg.

Huls, W. L., Brooks, D. L. and Bean-Knudsen, D. (1991). Response of adult New Zealand white rabbits to enrichment objects and paired housing. *Laboratory Animal Science 41*, 609-612.

Keenan, K. P., Ballam, G.C., Soper, K.A. et al. (1999). Diet, caloric restriction, and the rodent bioassay. *Toxicological Sciences 52*, 24-34.

Meller, A., Laine, O., Voipio, H-M. et al. (2004). Volatile organic compounds in animal bedding and enrichment items. Abstract. *9th FELASA Symposium Internationalization and Harmonization in Laboratory Animal Care and Use Issues* (14-17). Nantes, France, 2004.

Nevalainen, T. O., Nevalainen, J. I., Guhad, F. A. and Lang, C. M. (Submitted). Pair-housing of rabbits and variance in growth and serum chemistry.

Stauffacher, M., Peters, A., Jennings, M. et al. (2001). *Future principles for housing and care of laboratory rodents and rabbits*. Report for the revision of the Council of Europe Convention ETS 123 Appendix A for rodents and rabbits. Part B issued by the Council's group of experts for rodents and rabbits.

Whary, M., Peper, R., Borkowski, G. et al. (1993). The effects of group housing on the research use of the laboratory rabbit. *Laboratory Animals 27*, 330-341.

Correspondence to

Prof. Timo Nevalainen
NLAC / University of Kuopio
POBox 1627
70211 Kuopio, Finland
e-mail: Timo.Nevalainen@uku.fi

Importance and Effects of Enrichment on Physiology, Behaviour and Breeding Performance in Mice

Ping-Ping Tsai, Helge D. Stelzer, Anke Schraepler and Hansjoachim Hackbarth
Institute for Animal Welfare and Behaviour, School of Veterinary Medicine, D-Hannover

Summary
Sixty DBA/2 breeding pairs were used to understand the influence of enrichment on breeding performance. Four inbred strains of mice (A/J, BALB/c, C57BL/6 and DBA/2) were used for sequential experiments, in a total number of 240 (half for each sex), to collect information about the effect of enrichment on physiological traits and behavioural tests. Non-enriched and enriched groups were always provided with the same cage type to avoid the influence of floor space on agonistic behaviour, which can lead to different inter-individual distances.
According to the presented results environmental enrichment can affect experimental results and will not automatically improve well-being. Furthermore "Reduction" and "Refinement" may conflict with each other. Thus it will be necessary to evaluate the effects of environmental enrichment before an enrichment design is applied for a given experiment.

Keywords: mice, enrichment, variation, refinement and reduction

Introduction

In recent decades, there have been significant developments in laboratory animal science in the field of health monitoring (hygiene control), genetics and environment. This led to the possibility of standardisation in laboratory animal science. Reduction and refinement have been implemented in laboratory animal science in the last decades. Due to that the high level of standardisation has enhanced animal well-being by improving physiological health (refinement) and is largely responsible for the great decrease in the number of laboratory animals used in experimental research, due to the decrease in variation (reduction).

Currently environmental enrichment is intended for further improvement on the laboratory animal housing. Due to that a wide variety of enrichment designs are considered and the interactions between environment and genotype have been reported in many studies. The evaluation of enrichment has mostly been focused on the effects of experimental results (mean values) on the brain functions, behavioural performances and other parameters.

Concerning the reduction of 3Rs, the effects of enrichment on variation also need to be studied. Since the majority of statistical tests are basically comparing the size of the effect (the biological "signal") relative to the amount of variability in the data, the biological effects may be hidden by a large variation in an experiment. However, in contrast to the comparison of group means, only a few studies have aimed at the influence of enrichment on the variation (Eskola et al., 1999; Gärtner, 1999; Nevalainen, 1999; Mering, 2001; Tsai et al., 2002; Tsai et al., 2003b).

In the present study the response to enrichment was measured on the basis of breeding performance (for comparing the well-being status between different housing conditions), physiological variables (such as haematological data, body weight gain and relative organ weights), and behavioural tests (Open Field, Food Drive and Elevated Plus Maze) after long-term enrichment.

Non-enriched and enriched groups were always provided with the same cage type to avoid the influence of floor space on agonistic behaviour, which can lead to different inter-individual distances.

Material and methods

Housing: Two types of enrichment were chosen for the present study to evaluate the effects of enrichment. They contained: 1) a nest box, a wooden climbing bar and nest material according to Scharmann 1993 (E1) and 2) horizontal and vertical dividers, modified from Haemisch and Gärtner 1994 (E2). The purpose of the former was to meet the nest building and climbing behaviour of mice; the later was to offer the animals a burrowing system (see fig. 1 and 2).

Both control (non-enriched, NE) and test groups (enriched, E1 or E2) of each experiment were always provided the same cage size, as cage size can influence the agonistic behaviour and lead to different inter-individual distances.

Experiment 1: Following 4 weeks of adaptation 60 DBA/2 breeding pairs were randomly divided into three rack systems: a ventilated cabinet, a normal open rack and an individually ventilated cage rack (IVC rack) with enriched or non-enriched type II elongated Makrolon cages, half for each housing. Reproduction performance was recorded from 10 to 40 weeks of age for understanding the influence of enrichment (E1) on breeding performance (detail see Tsai et al., 2003a).

Experiment 2: A/J, BALB/c, C57BL/6J and DBA/2 were used for the subsequent experiments, in a total number of 240 (half for each sex). Animals at 3 weeks of age were marked and

assigned randomly to non-enriched (NE) or enriched (E1 or E2) type III Makrolon cages with equal numbers of cages in same-sex groups of four. Behavioural tests (Open Field, Food Drive and Elevated Plus Maze) were performed at 9, 10 and 11 weeks of age, respectively. At 14 weeks of age blood samples were collected for haematological analysis. The final body weights and organ weights were measured following euthanasia at 15 weeks of age to determine whether there are strain differences in the reaction to enrichment (detail see Tsai, 2002; Tsai et al., 2002 and Tsai et al., 2003b).

Statistic: The mean values were compared using factorial analysis of variance, followed by the Scheffé test (significance level 5%). To achieve independence from mean values, the coefficients of variation (SD/mean value, CV) were used instead of the variance (SD^2) or mean absolute deviation (MAD) to compare the variation between the non-enriched and enriched groups. As the CVs were not distributed normally, the CV of each variable was compared using the Wilcoxon signed rank test (nonparametric pair t-test).

Fig. 1: Enriched cage in experiment 1

Fig. 2: Enriched cage used in experiment 2 (E1 left, E2 right)

Results

Breeding performance

E1 housing (according to Scharmann 1993) did not improve the reproduction, but the variation of breeding performance increased due to enrichment (see tab. 1).

Physiological traits and behavioural tests

The effects of enrichment designs (E1 and E2) are not consistent, but vary according to the variables studied.

1. E1 housing had significant effects on Elevated Plus Maze performance, while significant differences were found in Open Field and Food Drive tests and in relative organ weights (adrenal, kidney, spleen and liver) due to E2 housing (detail see Tsai, 2002; Tsai et al., 2002 and Tsai et al., 2003b).
2. Strains reacted differently to enrichment (E1 and E2, detail see Tsai, 2002; Tsai et al., 2002 and Tsai et al., 2003b).
3. In comparison with NE groups there was a tendency towards an increased CV in enriched groups, especially in physiological traits and in Open Field and Food Drive tests (see tab. 2).

Conclusion

Our data showed that:

1. Environmental enrichment will not automatically improve well-being.
2. Enrichment may affect experimental results and can cause higher coefficients of variation (CV). Such influences were strain- and test-dependent.
3. The effects of enrichment on physiological traits are more focused on the variance than on the mean values, while enrichment significantly affected behavioural performance (group means and variations).

This indicates that the effects of enrichment designs vary according to strain and the variable studied. Furthermore reduction and refinement may conflict with each other. Thus, it will be necessary to evaluate the effects of environmental enrichment on variation, before an enrichment design is introduced into an experiment.

Tab. 1: Breeding performance of different housing conditions

	NE		E		Housing difference
	Mean	CV	Mean	CV	p
Total No. of litter/dam	5.44	37.9	4.15	57.7	0.0611s
Total No. of pups born/dam	24.9	51.5	15.9	77.1	0.0176s

CV = Coefficient of variation (SD/mean, %)
NE/E = 27 in non-enriched (NE) and 26 in enriched groups (E)
s = significant difference ($p<0.1$)

Tab. 2: The comparison of CVs between NE and E1/E2 groups

Strain	Physiological traits		Open Field and Food Drive		Elevated Plus Maze		Breeding performance
	NE vs. E1	NE vs. E2	NE vs. E1	NE vs. E2	NE vs. E1	NE vs. E2	NE vs. E1
DBA/2							4/7
A/J	14/22	13/22	3/4	2/4	2/12s	5/12	
BALB/c	8/22	14/22	2/8	4/8	2/12s	5/12	
C57BL/6	16/22s	11/22	6/8	7/8s	6/12	3/12s	
DBA/2	14/22	16/22s	7/8s	7/8s	8/12	6/12	
Total	52/88	54/88s	18/28s-	20/28s^{+}	18/48s	19/48	4/7

The CVs of each group were pooled and compared using Wilcoxon signed rank test (nonparametric test).
Data contained 4 strains of mice (A/J, BALB/c, C57BL/6 and DBA/2, both sexes), except A/J data of Food Drive test.
Physiological variables: relative body weight gain, relative organ weights and haematological data.
14/22 means: out of 22 compared variables, enriched groups had higher CV in 14 variables.
s: significant difference ($p<0.05$); s-: significant difference ($p<0.01$); s^{+}: significant difference ($p<0.001$).
For group means, E1 housing had significant effects on Elevated Plus Maze performance, while significant differences were found in Open Field and Food Drive tests and in relative organ weights due to E2 housing.

References

Eskola, S., Lauhikari, M., Voipio, H-M et al. (1999). Environmental enrichment may alter the number of rats needed to achieve statistical significance. *Scandinavian Journal of Animal Science 26(3)*, 134-144.

Gärtner, K. (1999). *Cage enrichment occasionally increases deviation of quantitative traits*. Proceedings of the International Joint Meeting 12th ICLAS General Assembly & Conference 7th FELASA Symposium. Palma de Mallorca, 207-210.

Haemisch, A. and Gärtner, K. (1994). The cage design affects intermale aggression in small groups of male laboratory mice: Strain specific consequences on social organisation, and endocrine activation in two inbred strains (DBA/2J and CBA/J). *Journal of Experimental Animal Science 36*, 101-116.

Mering, E., Kaliste-Korhonen, E. and Nevalainen, T. (2001). Estimates of appropriate number of rats: interaction with housing environment. *Lab. Anim. 35(1)*, 80-90.

Nevalainen, T., Lauhikari, M. and Eskola, S. (1999). Environmental enrichment may alter number of animals needed. International Joint Meeting XII ICLAS General Assembly & Conference / 7th FELASA Symposium, 51.

Scharmann, W. (1993). *Housing of mice in an enriched environment. Welfare and Science*, Proc. of the 5th FELASA Symposium, 335-337.

Tsai, P. P., Pachowsky, U., Stelzer, H. D. and Hackbarth, H. (2002). Impact of environmental enrichment in mice: 1. Effect of housing conditions on body weight, organ weights and haematology in different strains. *Lab. Anim. 36(4)*, 411-419.

Tsai, P. P. (2002). Impact of environmental enrichment and caging on the physiology and behavior of mice. Ph.D. Thesis. School of Veterinary Medicine Hannover.

Tsai, P. P., Oppermann, D., Stelzer, H. D. et al. (2003a). The effects of different rack systems on the breeding performance of DBA/2 mice. *Lab. Anim. 37(1)*, 44-53.

Tsai, P. P., Stelzer, H. D., Hedrich, H. J. and Hackbarth, H. (2003b). Are effects of different enrichment designs on the physiology and behavior of DBA/2 mice consistent? *Lab. Anim. 37(4)*, 314-327.

Correspondence to

Ping-Ping Tsai, PhD
Institut für Tierschutz und Verhalten
Tierärztliche Hochschule Hannover
Bünteweg 2
30559 Hannover
Germany
e-mail: tsai@tierschutzzentrum.de

Session 2.2
Pain, welfare and analgesia

Workshop of Experts: Definition, Recognition, Assessment, and Alleviation of Animal Distress in the Laboratory

Kathleen M. Conlee[1] *and Lesley A. Lambert*[2]
[1]The Humane Society of the United States, Animal Research Issues, Washington, DC, USA
[2]Linacre College, Oxford University, Oxford, UK

Summary
Distress in animal research is difficult to address and often overlooked, despite mandates in the U.S. to minimise animal distress. Consequently, The Humane Society of the U.S. held an expert workshop on animal distress, including whether creation of an operational definition of distress is possible. Discussion topics included what an operational definition of distress should encompass; prevention, causes, and measurement of distress; and regulation of distress. Participants agreed that creating a practical definition of distress is challenging, but that crafting a description of what might constitute animal distress, supported by specific examples, is possible. Workshop findings and future efforts on distress are discussed here.

Keywords: distress, animal welfare, refinement

Introduction

Distress in animal research is a major concern in regards animal welfare and quality of scientific results (Carstens and Moberg, 2000; Robertson, 2002), as well as public support of animal research (Aldhous et al., 1999; HSUS, 2001). Despite these issues and the fact that legislation in most countries requires minimisation of animal pain and distress, distress tends to be largely overlooked in the animal research laboratory. The main reasons for this lack of attention likely include the challenges of recognising, assessing and alleviating animal distress. Regardless of these challenges, animal distress can no longer be disregarded and must be made an urgent priority. As a result, The Humane Society of the United States (HSUS), through its Pain and Distress Campaign, has sought to increase attention to distress, particularly non-pain induced distress, by scientists, animal ethics and oversight committees, policymakers and the public. As part of this effort, The HSUS held a workshop of invited participants to discuss the state of knowledge of animal distress and whether it is possible to define distress in operational terms and to measure distress for application to animal research. The results of the workshop and future efforts to increase attention to distress will be discussed here.

Legislative requirements regarding distress in animal research in the United States

Animal Welfare Act

The Animal Welfare Act (AWA) and the Health Research Extension Act (HREA) are the two main laws that govern animal research in the United States. The AWA is enforced by the Animal Care division of the United States Department of Agriculture (USDA). The AWA pertains to warm-blooded animals *other than* purpose-bred mice, rats and birds, which make up an estimated 85-90% of warm-blooded animals used in animal research.

While the use of anaesthetics, analgesics, and tranquilising drugs was mentioned in the 1970 AWA amendments, it wasn't until 1985 that the AWA emphasised animal pain and distress. The 1985 amendments specify that

- Pain and distress are to be minimised;

- Anaesthetics, analgesics, and tranquilising drugs are to be used, unless there is scientific justification otherwise;
- Alternatives to procedures that cause pain and distress are to be considered by the investigator; and
- Each institution registered with the USDA must form at least one Institutional Animal Care and Use Committee (IACUC) to review animal use protocols and oversee the institution's animal care and use program.

Despite the fact that the term "distress" is used throughout the regulations, the USDA has never officially defined what is encompassed in its concept of distress. USDA Animal Care Policy #11 (Painful Procedures) defines "painful procedure" as "any procedure that would reasonably be expected to cause more than slight or momentary pain and/or distress in a human being to which that procedure is applied" (U.S. Department of Agriculture, 1997). The USDA policy defines "painful procedure" but not "distressful procedure", even though distress can be pain induced or non-pain induced. The absence of a regulatory definition of "distress" means that there is a lack of adequate guidance on distress. This could hamper or discourage institutions from expending the same effort in tackling distress as in addressing pain.

Health Research Extension Act

The Health Research Extension Act (HREA), passed in 1985, mandated that the National Institutes of Health (NIH) upgrade its requirements for animal research oversight. The HREA applies to all research facilities that receive federal funds from the NIH or its parent agency, the Public Health Service (PHS). The provisions of the HREA were implemented through the PHS "Policy on the Humane Care and Use of Laboratory Animals", which is enforced by the NIH's Office of Laboratory Animal Welfare (OLAW) and applies to all vertebrate species, thereby partly compensating for the exclusion of birds, mice, and rats from the AWA.

PHS Policy, like the AWA, emphasises avoidance or minimisation of pain and distress and calls for the formation of an Institutional Animal Care and Use Committee to review animal protocols and oversee the institution's animal care and use program. The "U.S. Government Principles for the Utilisation and Care of Vertebrate Animals Used in Testing, Research and Training" are also incorporated into PHS Policy. Three of the nine government principles directly address animal distress and pain, demonstrating the importance of these issues. Finally, PHS Policy calls upon research facilities to follow the provisions in the publication the "Guide for the Care and Use of Laboratory Animals" (often referred to as The Guide). The Guide recommends consideration of alternatives for all animal-based research, emphasises the importance of minimising distress, and offers examples of procedures that have the potential to cause distress and pain, such as physical restraint, multiple major survival surgery, food or fluid restriction, and the use of abnormal environmental conditions (Institute for Laboratory Animal Research, 1996). The Guide, however, does not clarify what is encompassed in its concept of distress, despite the fact that the term is used throughout the text.

Recent regulatory efforts

On July 10, 2000, the USDA published a call for public comment regarding whether the term "distress" should be defined and, if so, what the elements of the definition should be; and whether the current system used by research institutions to classify and report pain and distress should be improved. The published call for comments emphasised the advantages of these efforts, including helping research facilities to recognise and minimise animal distress in accordance with the AWA.

There were over 2,800 comments submitted in response to USDA's call for public comments, approximately 2,150 of which were from the research community. The HSUS examined all of the comments submitted by the research community and found that even though a majority of these respondents opposed the adoption of a definition of distress, 97% supported a definition developed by the National Research Council (NRC)[1] should the USDA decide to adopt a definition. In other words, although the research community doesn't want the USDA to define distress or otherwise expand its authority over research practices (a standard response when enhancements of the AWA are considered), they nearly all agree on the definition that should be used if the USDA does decide to move forward.

As of March 2006, the USDA has taken no official action on defining the term distress or changing the pain and distress categorisation system. The HSUS, however, believes that the adoption of a definition of distress and a new pain and distress categorisation system will increase attention to distress, will provide guidance on tackling distress, and will prompt needed data collection and research on distress.

Distress workshop convened by The HSUS

The HSUS held a workshop to discuss the state of knowledge of animal distress and whether it is possible to define distress in operational terms and to measure distress for application to animal research. The workshop was held February 11-13, 2004 in Baltimore, Maryland. Seventeen international participants were invited from diverse disciplines, representing the fields of laboratory animal medicine, animal physiology, animal behaviour and applied ethology, animal welfare, ethics and philosophy, regulation of animal research, and animal protection.

Workshop design, key objectives and discussion topics

The workshop was designed in order to meet pre-determined objectives. The key elements of the workshop design included:

1. Bringing together stakeholders from a range of disciplines;
2. Focusing on a manageable number of specific topics;
3. Examining and attempting to integrate a range of animal welfare-related scientific findings, including production,

[1] The definition of distress published by the National Research Council (Institute for Laboratory Animal Research 1992): "an aversive state in which an animal is unable to adapt completely to stressors and the resulting stress, and shows maladaptive behaviors."

behaviour, physiology, and veterinary health;
4. Identifying areas of consensus and determining how to communicate findings to a relevant audience;
5. Identifying barriers to consensus; and
6. Clarifying requirements for future efforts on specific issues of definition, measurement, and alleviation of laboratory animal distress.

A number of efforts were made in order to stimulate discussion at the workshop, such as advanced preparation of briefing papers by some participants, consideration of specific questions and sharing of answers with all participants, and brief presentations by some participants at the workshop. Importantly, the workshop was structured to allow for lengthy discussion of specifically designated topics; points of contention and agreement were recorded.

The key objectives of the workshop were to:
1. Present the current state of knowledge on the definition of distress and science of distress recognition, measurement, and alleviation from a range of disciplines;
2. Reach a working agreement regarding an operational definition of distress, signs of distress, and modes of measurement;
3. Clarify requirements for future efforts on the definition, measurement, and alleviation of laboratory animal distress; and
4. Publish the workshop proceedings via print or electronic media.

With these objectives in mind, the topics posed for discussion at the workshop included:
- determination of what should be incorporated into the concept of distress;
- the reasoning behind incorporating these factors into the concept of distress;
- causes of distress and how this information can ultimately be used for prevention; measurement, scoring and validation of distress;
- alleviation of distress;
- incorporating an understanding of distress into regulation,
- finally but importantly, practical application of information to the laboratory.

Areas of general agreement: universal

General points of agreement were recorded, discussed and finalised at the workshop. Some of the points will be discussed briefly here, but additional points and more in-depth discussion of these issues will be provided in a future published executive summary of the workshop authored by workshop participants.

A general point that all participants agreed upon is that distress is of most concern when it is more than momentary or slight, and those are cases that should be priority in terms of attention. Extent of distress can also vary according to duration and intensity of the stimulus and in the context in which it occurs. An animal's distressful experience may represent states similar to (but not necessarily exactly the same as) negative human states, such as anxiety, boredom, hunger and malaise, for example. Animals may also experience states of distress that we simply aren't able to identify as of yet.

The terms "pain and distress" are often seen in conjunction, particularly in regulations and legislation, but it must be emphasised that distress is not necessarily related to, nor a consequence of, pain. For example, an animal may be anxious to the point of distress, but anxiety is not caused by pain. Distress is also difficult to define and has different meanings in different contexts. For example, distress can mean one thing to regulators, another to the public, and yet another to scientists working with animals in research. The experts agreed that in order to deal with the difficulties of defining distress and the issue of different contexts, a broad working description could be created and accompanied by specific examples of what would cause different levels of distress (such as mild, moderate and severe).

Areas of general agreement: causation, recognition and assessment of distress

There was lengthy and highly productive discussion of causation, recognition and assessment of distress. As with other topics, general areas of agreement were found. There was unanimous agreement that distress can arise from various sources, such as those related to the experiment, housing and husbandry, as well as genetic factors.

In attempting to determine when an animal could be experiencing distress, the idea of using two broad questions was discussed (Dawkins, 2004):
1. Is the animal healthy?
2. Does the animal have what s/he wants?

There are a number of resources that address how to determine the health of an animal (such as courses provided at www.researchtraining.org); therefore the means of answering that question will not be discussed here. In terms of considering whether an animal has what s/he "wants", in general, "wants" are considered resources or behaviours that are important to the animals (Dawkins, 2004), without which the animal may experience unpleasant psychological states and may consequently develop or behave abnormally. The importance an animal places on gaining a resource, or removing itself from a particular situation, provides a basis on which to determine how distressed the animal might be without this resource.

Both physiological and behavioural measures can provide information about an animal's health and psychological state. However, there is no single measure that is sufficient to indicate distress in any given situation; therefore, a selection of appropriate measures should be considered when assessing distress. The type of measures used to assess distress are often complex and the interpretation of results may be guided by looking at the species evolutionary history. Ongoing daily experience, familiarity with individual animals, knowledge of species-typical behaviour, and high quality diagnostic skills also play an important role in recognising distress. Experienced clinical observation and divergences from the animal's normal state were also identified as key to identifying potential acute distress.

Aside from addressing distress, there was a brief discussion surrounding how to improve well-being so that the animals have positive experiences. This brought up a third question that could be asked, namely, how can we determine what resources the animal requires in order to perform behaviours and social interactions that provide positive experiences? It was agreed that those working with animals in research should maximise opportunities for animals to perform behaviours that are important to them.

Areas of general agreement: Recommendations for oversight

The participants recognised that oversight is an important factor in addressing distress in animal research. Three main recommendations were made in regards to oversight: a team approach must be taken, there should be post-approval protocol monitoring, areas of needed research should be determined, and guidance from professional bodies is currently lacking, yet needed.

Areas of agreement: development of resources for oversight bodies and researchers

If there is going to be increased attention to, recognition of, and prevention of distress, the research community urgently needs additional resources. Some resources that should be taken into consideration include a journal that focuses on refinement; increased information in current journals regarding how pain and distress were minimised, such as how humane endpoints were determined; identification of research priorities related to distress by those qualified to do so; earmarked funding for the set priorities; and, finally, development of good practice guidelines, written by experts.

Future efforts to increase attention to animal distress

Several products have been developed from this workshop. An executive summary of the workshop authored by participants has been submitted to a laboratory animal science journal for publication. Furthermore, The HSUS is coordinating publication of a technical book on distress and pain entitled "Recognition and Alleviation of Pain and Distress in Animal Research", authored by international experts and to be published by Humane Society Press in 2006. The intended audience of the publication encompasses laboratory animal veterinarians, scientists, members of Institutional Animal Care and Use Committees (or comparable bodies outside of the United States), caretakers, and other laboratory personnel, as well as regulatory authorities and policymakers.

The purpose of the book is to promote animal well-being in terms of operational consideration and management of animal distress and pain, draw attention to the latest developments in recognition and alleviation of pain and distress in animal research and testing, and to discuss related policy issues. A number of topics concerning pain and distress in animal research and testing will be included, such as philosophical aspects and concepts; international regulatory definitions of distress; causes and measurement of distress; distress in the context of pain; understanding stress, distress and suffering; animal illness and distress; animal emotions, psychological well-being and distress; the role of clinical veterinary medicine in assessment and treatment of distress, and resolving animal distress (with specific examples of good practice provided). Practical application to the animal research laboratory is what will make this publication on pain and distress unique.

Conclusion

Although there has been little work on laboratory animal distress at a regulatory level in the United States for some years, this workshop demonstrated that the issue of distress is neither dormant nor dominated by dissenting scientific voices. The workshop participants, all experts in their respective scientific fields, reached important general agreement on the contexts in which distress might arise and the types of negative experiences that might constitute distress. The lively debate produced a range of proposals and practical solutions for moving forward the discussion about defining distress, as well as methods for recognition, assessment, alleviation and prevention of distress. The forthcoming book, developed directly as a consequence of the workshop, will assist those making difficult operational and ethical decisions in the laboratory to make informed, comprehensive judgments on animal distress. It is hoped that this workshop and its associated publications will provide further stimulus toward attention to animal distress, to benefit both animals and science.

References

Aldhous, P., Coghlan, A. and Copley, J. (1999, May 22). Let the people speak. *New Scientist*, 26-31.

Carstens, E. and Moberg, G. P. (2000). Recognizing pain and distress in laboratory animals. *ILAR Journal, 41*, 62-70.

Dawkins, M. S. (2004). Using Behaviour to Assess Welfare. *Animal Welfare 13 (Supplement)*, S3-S7.

HSUS (The Humane Society of the United States) (2001). Opinion Poll on pain and distress in research. Retrieved October 14, 2005 from http://www.hsus.org/animals_in_research/pain_distress/opinion_poll_on_pain_and_distress_in_research.html.

Institute for Laboratory Animal Research (ILAR) (1992). *Recognition and Alleviation of Pain and distress in Laboratory Animals*. Washington, DC: National Academy Press.

Institute for Laboratory Animal Research (ILAR) (1996). *Guide to the Care and Use of Laboratory Animals*. Washington, DC: National Academy Press.

Robertson, S. A. (2002). Pain management in laboratory animals – are we meeting the challenge? JAVMA, 221, 205-208.

U.S. Department of Agriculture (1997). Painful procedures. Retrieved October 14, 2005 from http://www.aphis.usda.gov/ac/policy/policy11.pdf.

Correspondence to

Kathleen M. Conlee
Director of Program Management
The Humane Society of the United States
Animal Research Issues
Washington, DC, 20037, USA
e-mail: kconlee@hsus.org

Dr. Lesley A. Lambert
Linacre College, Oxford University
St. Cross Road
Oxford, OX1 3JA, UK
e-mail: Lesley.king@linacre.ox.ac.uk

Use of Analgesics in Experiments

Martin Sager
Universitätsklinikum, Düsseldorf, Germany

Summary
Refinement in animal experiments means to diminish pain, suffering and harm. The use of analgesics in experiments is applied refinement and an essential part of good veterinary care.
It was the task of this report to give a comprehensive overview about the topic of pain treatment in laboratory animals. This report is based to the manual of the working group for anaesthesia and analgesia of the German Society for Laboratory Animal Science (GV-SOLAS) published in the year 2002. By the extent of this report it is only possible to give proposals by example for pain relief in different species used in the diverse kinds of experiments.

Keywords: analgesics, pain relief, postoperative care

Introduction

Profound knowledge is needed to rule out the most effective pain management protocol by use of general anaesthesia, regional anaesthesia and systemic analgesics. This has to take into account the side effects of drugs used to physiological parameters, behaviour and pathophysiological patterns.

The concept of the cited recommendations is explained, which provides easy access to the necessary information to investigators.

This paper should help researchers to plan the animal experiments with concern for an adequate pain treatment protocol, which provides a maximum of benefit for the animals as well as optimal conditions for a high scientific output.

Representatives of the authorities working in the field of animal welfare giving license or controlling animal experiments should find very easily guidelines and diverse alternatives for pain relief in animals.

The paper should be also a formula for all people caring for animal welfare.

The influence of pain to animal experiments – why alleviate pain? (Pathophysiological aspects)

There is a relevance of pain therapy for laboratory animals first by ethical aspects and second due to most national animal welfare acts. Therefore, adequate pain therapy needs to be carried out in any animal experiment where post-operative pain occurs.

Furthermore, we have to point out that there is an effect of pain to the results of animal experiments. The laboratory animal is not a passive measuring instrument in the experiment.

Pain provokes pathophysiological effects which can influence experimental results, more than the use of pain killers *per se*. Long duration and intensity of pain can cause death.

Pain has pathophysiological effects on different organ systems as listed below (Jage, 1997; Larsen, 2002; Henke and Erhart, 2001).

Endocrine System
- Increased secretion: catecholamines, corticoids, glucagon, growth hormones, ACTH, ADH
- Decreased secretion: insulin, testosterone
- Effect: adynamia in mobilisation, floppiness, amyotrophia, extended anastasis
- Fluid retention, oliguria, disbalance of electrolytes
- High metabolic rate with increased oxygen consumption
- Effects on organ systems

Neuroendocrine system
- Increased β-endorphin index
- Blockade of NMDA-channels

Sympathetic nervous system
- Increased activity, release of catecholamine during excess stimulation of the sympathetic nervous system
- Decreased tissue perfusion with increased tissue acidosis, resulting in hyperalgesia
- Risk of hypoxia in poorly perfused organs (heart, brain, intestine, lung)
- Atonia of the gastrointestinal tract, risk of paralytic ileus
- Activation of the renin-angiotensin-aldosterone system during a compromised renal perfusion
- Thereby inducing peripheral vasoconstriction
- Increasing platelet aggregation
- Increasing release of noradrenalin from peripheral nerves causing hyperalgesia

Immune system
- Immune suppression (susceptibility to infection)
- Impaired wound healing

Tab. 1: Synopsis of the effects and side effects of the most known agents

	Active ingredient	Trade name	Application area	Side effects	Comment
Opioids	Piritramide	Dipidolor	Strong pain of all kinds	Long-time treatment (>3 day) because of constipation	Titrate with strength monitoring of the respiration
	Fentanyl-plaster	Durogesic	Evenly strength analgesics about 3-5 d by dogs, rabbits, pigs, ruminant animals	Rodent, immediate analgesics, fever, respiratory depression	Beginning therapy hole for 6-12h, good bandage needed
	Tramadol	Tramal	Weaker pain, continuous drop infusion (CDI)	Strong pain	Very short effect, possibly CDI
	Pethidine	Dolantin	Weaker pain, on bowel, A-bile and pancreas duct due to spasmolysis	Long-time application	Very acute effect
	Buprenorphine	Temegesic	Medium- string pain, for part ial antagonism of effects of μ-agonists	A-bile and pancreas duct examination, constipation by application >5d	In low doses excitatory, in high doses sedation, allotriophagy, CAVE by preoperative. administration
	Butorphanol	Morphasol	Temporary therapy low - medium pain	Long-time treatment, strong pain	Excellent antitussive
Antipyretics	Acetylsalicyl acid	Aspirin, ASA	Low inflammatory pain, accessorily an antithrombotic effect, fever	Bleeding risk, GIT average	CAVE not using it pre op. because of bleeding risk
	Metamizole	Vetalgin, Novalgin	Low-medium pain, especially in abdomen, spasmolysis, long-time treatment just as CDI	Long-time treatment	Good in combination with opioids CAVE very slow i.v.
NSAIDs	Carprofen	Rimadyl	All kinds of inflammatory pain, especially by operation, also for acute and chronic inflammation of the musculoskeletal system	Gastrointestinal disorder	Good effects, when you use it for the insult, long range treatment, good effects in combination with opioids
	Flunixin-Meglumine	Finadyne	Acute and chronic inflammatory pain, endotoxin shock		Nephrotoxic
	Ketoprofen	Romefen	See above, addition on eyes		
	Meloxicam	Metacam	Pain states all kind		Good effects, when you applicate it for the insult, long range applicable, good effects in combination with opioids
	Phenylbutazone	Tomanol, Phenylarthrit	Anti-inflammatory, especially the musculoskeletal system, contingent ruminant animal		Many side effects, not good for routine use
	Tolfenamic acid	Tolfedine	Acute and chronic inflammatory pain, fever loweing		Especially good for cats

- Inhibition of the mitotic rate and locomotion of T-cells
- Inhibition of the lymphokine production, inhibition of phagocytosis
- Decrease in interleukin release, cell immunity, tumor immunity, host defence, production of antibodies

Blood
- Depletion in the spleen and skin: Lymphopenia, eosinopenia, neutrophilia

Respiratory system
- Especially affected after thoracic and abdominal intervention
- Reduced tidal volume and vital capacity, laboured breathing
- Symptoms: atelectasis with disturbances of the pulmonary gas exchange, therefore resulting in the development of infection, pneumonia
- Respiratory and metabolic acidosis

Cardiovascular system
- Occurs after the activation of the sympathetic system: tachycardia, peripheral vasoconstriction with increased vessel resistance, increased heart contractility and myocardial O_2-consumption and hypertension.

Gastrointestinal system
Cause: Excitation of peritoneal nerves, increased sympathetic activity and ischemia
- Symptoms: gastrointestinal atonia, risk of paralytic ileus, sickness, vomiting, distension of the intestine and higher abdominal pressure, reduced diaphragm movement with pulmonary dysfunction
- Excitation of visceral nerves
- Disturbance in visceral perfusion causing ischemia, thereby aggravating pain, reduced food and water intake (hypoglycaemia, dehydration)

Urogenital system
- Decreased motility in the urinary tract
- Urine retention

Musculature
- Spasms, tremor, convulsions
- For longer periods: adynamia, floppiness, amyotrophia

Behaviour
- Depression, hyperagressiveness, self-mutilation
- reduced grooming activity (especially in rodents)

→ Behavioural changes induced by pain in animals (particularly small ones) are a good indication for the evaluation of severity

Systemic analgesics (for p. op. supply)

Advisable are (see tab.1):
- Opioids (act predominantly central):
Buprenorphin, Piritramid, Pethidin, Butorphanol, Tramadol and Fentanyl-Patch.
- Antipyretics (act central and peripheral):
Acetylsalicylic acid, Metamizol.
- NSAIDs (act predominantly peripheral):
Carprofen, Etodolac, Flunixin-Meglumin, Ketoprofen, Meclofenamin acid, Meloxicam, Niflumin acid, Phenylbutazon (ruminant), Piroxicam, Tepoxalin, Tolfenamin acid, Vedaprofen.

Application in form of: Plaster, suppositories, SC, PO, IV, CID, osmotic mini pumps. It is not advisable to use uncritically human analgesics like Diclofenac and Paracetamol in experimental animals (toxic, to short a half-life period).

The following particulars about the dosage table are based on appropriate literature and the experience of the authors (see tab. 2 and 3). The dosage must be adapted to the respective clinical situation and the aim of the experiment. Frequent side effects will be found in the information letters of the product, but most apparent side effects in a routine use will be found in the following table. The acute appearance of side effects is often based on a to rapid i.v. injection. Acute side effects are close to a certainty if intravenous injections are performed too rapidly. The dosage plans are given exemplary for guinea pig, chinchilla, mouse, hamster and dog. For recommendations for other species see the recommendation of GV-SOLAS (2002).

Tab. 2: Dosage table for the guinea pig, chinchilla, mouse and hamster

Substance	Dose (mg/kg)	Administration	Interval of administration	Notes
Buprenorphine	0.05-0.1	s.c./i.p.	8h	not hamsters
Butorphanol	1.0-5.0	s.c.	4-6h	not hamsters
Flunixin-Meglumine	3.0-5.0 (mouse)	s.c.	12h	
ASA	120.0-300.0	p.o.	24h	not hamsters, guinea pigs
Metamizole	80.0(guinea pig) (*) supplementary stress see Thoracoscope, Laparoscope – 2 drops/animal) 200.0 (mouse) 100.0 (hamster) (0.5-2 drops of 1:4 diluted solution/animal)	p.o.	every 6h	
Carprofen	4.0-5.0	s.c.	24h	

Tab. 3: Dosage table for the dog

Substance	Dose (mg/kg)	Administration	Interval of application	Notes
Buprenorphine	0.01-0.02	i.v., i.m., s.c.	8-12h	
Butorphanol	0.1-0.5-1.0	i.v., i.m., s.c.	1-2h	
Fentanyl-Patch	5-10kg: 25 mg/h 10-20kg: 50 mg/h 20-30kg: 75 mg/h from 30kg:100 mg/h	transdermal	48-72h	CAVE oral intake
Pethidine	2.0-6.0	i.m., s.c.	1-2h	spasmolytic at musculature
Piritramide	0.1 0.2 0.1 (-0.3)/h	i.v. s.c. i.v.	1-2h 2h DTI	
Tramadol	1 (-3)/h	i.v.	DTI	
ASS	25.0 10.0	p.o. slowly i.v.	6-8h	
Metamizole	20.0-50.0	slowly i.v.	4h	spasmolytic at musculature
Carprofen	4.0 or 2.0	i.v., i.m., p.o., s.c.	24h 12h	
Flunixin-Meglumine	0.5-1.0	i.v., i.m., s.c., p.o.	24h	over max. 3d at endotoxin-shock every 12h
Ketoprofen	1.0-2.0 from 2.d: 0.5-1.0	i.v., i.m., s.c., p.o.	24h	over 3-5d
Meloxicam	0.1	p.o., s.c., i.v.	24h	
Phenylbutazone	5.0-10.0 20.0-60.0/d	slowly i.v., i.m., p.o.	24h	then reduce, max. 800/animal/d
Piroxicam	0.3	p.o.	48h	
Tolfenamic acid	4.0	s.c., p.o.	24h	max. 3d

Recommendation for analgesic methods in experimental procedures

The cited report gives also recommendations for the alleviation of pain in selected experimental standard procedures. These procedures are listed in a topographical synopsis. For examples see tables 4 to 7.

Pain dispensary

In general: For postoperative analgesia it is enough for the routine use of an experimental unit to have one substance from the different substance groups (opioides, antipyretics, NSAIDs, local analgesics, analgesic anaesthetics). It will be impossible to work without opioide substances. With these analgesics you can treat nearly all species – for example:

Stocking up of postoperative analgesics:

Examples: Buprenorphin, Metamizole, Carprofen (standing in for NSAIDs of the new generation like also Meloxicam, Flunixin-Meglumine, Tepoxalin), Bupivacaine, Ketamine, Methadone (for the after-effect in the postoperative phase). Other substances could be needed after experiment planning.

In my experience you can treat nearly 95% of all studies just with Buprenorphin, Metamizol or Carprofen (or in combination) to get enough analgesic effect (tab. 8).

Dosage instructions

For all animal species and all used substances it is recommended to make a dosage instruction (tab. 9).

Remark to legal regulations

Veterinary bodies in the European Union and in Germany are confronted with a lot of legal regulations to deal with opioide pharmaceutical products.

Conclusions

Analgesic treatment should be the standard in all not insignificantly painful experimental procedures. It was pointed out that less stressful and painful protocols lead to better and more valid results. Indeed, the use of analgesics is not recommended in all experiments because effects and side-effects of the drug may interfere with the expected results. But painful procedures without any analgesic treatment should be limited to the absolutely necessary exception.

The data presented should be a base for an intensive discussion by all persons involved in designing better animal experiments.

References

Bonica, J. J. (1985). Treatment of Cancer Pain: Current Status and Future Needs. In H. L. Fields, R. Dubner and F. Cervero (eds.), *Advances in Pain Research and Therapy (589, Vol. 9), Proceedings of the Fourth World Congress on Pain*. New York: Raven Press.

Bonica, J. J. (1990). Cancer Pain. In J. J. Bonica (ed.), *The*

Tab. 4: Recommendations for the alleviation of pain in selected experimental procedures

Facial cranium						
Tissue/Organ	**Route**	**Example**	**Extent of pain stress (intensity and duration of pain)**	**Duration of analgesia**	**Animal species**	**Suggestions for therapy**
Oral cavity	oral	Stomatitis, Gingivitis; intervention in teeth; jaw-bone: implants, partial resection	dog low-grade, cat moderately till high-grade low-grade, Pulpitis moderately at adequate stability low, instability and at injuries of *N. alveolaris* moderately till high-grade disease and damage by anorexia	3-5d	dog, pig	antiphlogistics (steroidal and non-steroidal) at harms of cranial nerves often opioids necessary
Maxillary articulation	various	implants, resection	dependent on the masticating exposure myofacial syndrome of masticating muscle symptoms with convergence to other structures in the service area of *N. trigeminus* moderately till high-grade disease and harms by anorexia	3-5d	dog, pig	NSAIDs take care of muscle relaxation sometimes opioids necessary
Paranasal sinus, (frontal) sinus		tumour development	headache similar low-grade till moderately		mouse, rat	opioids
Eye, Eye socket		all ophthalmologic models	irritations of cornea or of N. opticus, raised intraoculare pressure extrem painful	3-5d	all	antiphlogistics (steroidal and non-steroidal cornea local glaucoma therapy if required opioids
Inner, middle ear		bulla osteotomy	very sensitive structures (*N.facialis, N. trigeminus*)	3-5d	all	NSAIDs most with opioids
External auditory canal			only if the duration of injury is very short you can do without analgesic			
Horn		amputation	highly innervated	1-3d	sheep	Effective: local anaesthesia, NSAIDs

Tab. 5: Recommendations for the alleviation of pain in selected experimental procedures

Neurocranium						
Tissue/Organ	**Route**	**Approach**	**Degree of pain (intensity and duration)**	**Duration of analgesia**	**Animal species**	**Suggestions for therapy**
Neurocranium	trepanation, bore hole	stereotactic operation probe implantation inoculation of tissue, cells or infectious substances tumour implantation	painful at aditus and irritations p. op. (instability of implants, infections) periosteum and meninges are very pain sensitive direct p. op. low at increase of intracranial pressure painful (tumour growth, oedema), emission to facial and cervical region epileptic seizure could increase the disease	1-3d p. op. 0-1d at existing tumour permanent therapy no-go criteria	all mouse rat	opioids in combination with steroidal antiphlogistics if low: Metamizole, NSAIDs treat the oedema (Mannit-infusion) anticonvulsants, sedatives
		general brain trauma	at increase of intra cranial pressure painful	3d	mouse rat	treat the oedema (Mannit-infusion) Metamizol NSAIDs

**Tab. 6: Recommendations for the alleviation of pain in selected experimental procedures.
Pain in neoplasia.**

Lesions of the tissues (Stumpf, 1993)	**Examples**	**Extent of pain load (intensity and duration of pain)**	**Duration of treatment**	**Com. used species**	**Recommendations**
Tumor growth: Bone- and soft tissue infiltration Compression and infiltration of nerve-, blood- and lymph vessels Edema with impaired perfusion Skin: necrosis, ulceration and secondary infection **Caused by therapy:** Radiation: - fibrosis, neuropathy, - radiation osteomyelitis, - Chemotherapy: - inflammation, neuropathy	Cancer models	**Non-uniform:** Bony and periosteal pain, soft tissue pain, radicular pain, loss of function, irritations Primary by cancer 60-90%, by therapy about 5% of pain load (Twycross and Fairfield, 1982) 37% already in an early stage painful (human)! **Pain prevalence:** (Zech et al., 1988; Bonica, 1985) 60% soft tissue-, 75-80% bone cancer usually severe (Bonica, 1990)	Longtime-therapy	Mice, rats	steroidale and not steroidale antiphlogistics, Metamizol, in advanced stage need of complementary opioides use of local anaesthetic techniques! endpoints!

Tab. 7: Recommendations for the alleviation of pain in selected experimental procedures

Transgenic technology						
Model	**Aditus**	**Example**	**Intensity and duration of pain**	**Duration of the analgetic therapy**	**Frequent animal species**	**Suggestions for therapy**
Vasectomy	Abdomen, ventral	Creation of infertile males for preparation of pseudogravid wet nurses	Low till median pain stress for 1 day	1 day	Mouse, rat	NSADIs, Metamizol, Opioids
Epidydectomy	Skrotum	Creation of infertil males for preparation of pseudogravid wet nurses	Low till median pain stress for 1 day	1 day	Mouse, rat	NSADIs, Metamizol, Opioids
Embryo-transfer	Abdomen, dorsolateral	Hygienic clean up; transmission of embryos after cryoconservation, prenucleusinjection, blastcystinjection, *in vitro* fertilisation (IVF), intracyto-plasmatic spermatozoainjection (ICSI)	Low till median pain stress for 1 day	1 day	Mouse, rat	NSADIs, Metamizol, Opioids
Implantation of ovary	Abdomen, dorsolateral	Preservation of mouselines through transplantation of ovaries	Low till median pain stress for 1 day	1 day	Mouse, rat	NSADIs, Metamizol, Opioids
Biopsy to get DNA	Tag, auricle, hair follicle, oral cavity, terminal bowel		Low pain tension under 1 day, no analgesia needed		Mouse, rat	
Mark	Ear crenation\ punching\ mark, transponder, tattoo	Identification	Low pain tension under 1 day, no analgesia needed		Mouse, rat	

Tab. 8: Main indication and -contraindication to use Buprenorphin, Carprofen and Metamizol

	Buprenorphin	**Carprofen**	**Metamizol**
to recommend	Middling till strong achestate each kind	Each inflammation ache also with injury and operation, especial in prevention	Not by inflammation, particularly use by cramps and spasm of the hollow entrails
inadvisable	Exploration of GIT-motality and bile secretion	Exploration in context of inflammatory	Animals who are to often been stressed by (i.m.) application and cats, because they could salivate to much

Tab. 9: Pain therapy with Buprenorphin, Carprofen and Metamizol

	Buprenorphin	Carprofen	Metamizol
Goat	0.006 i.v/ i.m./ s.c. 12h	4 i.v/ i.m. /s.c/ p.o. 12-24h	20-50 i.v/ i.m. 4h
Pig	0.05-0.1 i.v/ i.m./ s.c. 8-12 h	4 i.v/ i.m./ s.c./ p.o. 12-24h	20-50 i.v/ i.m. 4h
Dog	0.01-0.02 i.v/ i.m./ s.c. 8-12 h	4 i.v/ i.m./ s.c./ p.o. 12-24h	20-50 i.v/ i.m. 4h
Cat	0.005-0.01 i.v/ i.m./ s.c. 8-12 h	4 i.v/ i.m./ s.c./ p.o. 24h	20-50 i.v/ i.m. 4h
Rabbit	0.0075 i.v/ i.m./ s.c. 6-12 h	2-4 i.v/ i.m./ s.c./ p.o. 12-24h	20-50 i.v/ i.m. 4h 3-5 drops p.o 4h
Guinea pig	0.05- 0.5 s.c. 6-12h	4 i.m./ s.c. 12h	80 (1/2-2 drops/animal)p.o. 4h
Rat	0.05- 0.5 s.c. 6-12h	5 i.v/ s.c.. 12h	110 (2 drops/animal) p.o. 4h
Gerbil		4 s.c. 24h	100 (1-2 drops/animal) p.o. 4h
Hamster		4 s.c. 24h	100 (1-2 drops/animal) p.o. 4h
Mouse	0.05- 0.1, s.c. 6-8h	5 i.v./ s.c. 12h	200 (1-2 drops/animal) p.o. 4h

Management of Pain (400). Philadelphia: Lea & Febiger.

Henke, J. and Erhart, W. (2001). *Schmerzmanagement beim Klein- und Heimtier*. Stuttgart: Enke.

Jage, J. (1997). *Schmerz nach Operationen*. Stuttgart: Wissenschaftliche Verlagsgesellschaft mbH.

Larsen, R. (2002). Postoperative Schmerztherapie. In R. Larsen (ed.), *Anästhesie* (783-819). München: Urban & Schwarzenberg.

Report of the Working Group for Anaesthesia and Analgesia of the GV-SOLAS (2002). Pain therapy for laboratory animals. Publication of GV-SOLAS, htpp://www.gv-solas.de.

Strumpf, M. (1993). Krebsschmerz. In M. Zenz und I. Jurna (Hrsg.), *Lehrbuch der Schmerztherapie* (715-726). Stuttgart: Wiss. Verl.-Ges.

Twycross, R. G. and Fairfield, S. (1982). Pain in far-advanced cancer. *Pain 14(3)*, 303-310.

Zech, D., Schug, A. and Horsch, M. (1988). *Therapie-kompendium Tumorschmerz*. Erlangen: Perimed.

Acknowledgements

I like to acknowledge my co-workers in the working group for anaesthesia and analgesia of the GV-SOLAS: M. Arras, K. Becker, W. Erhardt, J. Henke and B. Kuhnt.

Correspondence to

Dr. med. vet. M. Sager
Universitätsklinikum Düsseldorf
Moorenstrasse 5, Geb.22.22
40225 Düsseldorf, Germany
e-mail: sagerm@uni-duesseldorf.de

Session 2.3
Non-invasive approaches – new imaging and remote techniques (Doerenkamp-Zbinden Session)

Magnetic Resonance Imaging of Animal Brain *In Vivo*

Jens Frahm, Susann Boretius, Takashi Watanabe and Thomas Michaelis
Biomedizinische NMR Forschungs GmbH, Max-Planck-Institut für biophysikalische Chemie, Göttingen, Germany

Summary
As a non-invasive tool for mapping anatomical and functional aspects of the central nervous system, magnetic resonance imaging (MRI) reduces the number of animals in follow-up examinations and refines the quality of the results by yielding in vivo data of individual animals over extended time periods. Although applications cover the full range from non-human primates to insects, the vast majority of MRI studies focuses on rodents. This work illustrates selected applications in mice based on three-dimensional images at almost 100 µm isotropic spatial resolution. A primary aim is a better understanding of the pathophysiological mechanisms underlying human brain disorders. In this respect, the cuprizone mouse model allows the identification of MRI markers for demyelination and remyelination. Complementary, the use of transgenic mice with a Cnp1 deficiency in oligodendrocytes adds selective MRI studies of axonal damage without demyelination. A second goal is the functional assessment of synaptic activity and axonal transport in various brain systems using manganese-enhanced MRI. The contrast agent identifies activity-dependent regional differences in the mouse hippocampus and delineates the efferent pathways to the ventral hippocampal commissure and septal region. While opening the way for investigations of memory and learning in mutant mice, corresponding studies of the habenulo-interpeduncular system provide links to cognition and behaviour and the mechanisms of neuropsychiatric diseases.

Keywords: axonal damage, axonal transport, cuprizone, demyelination, magnetic resonance imaging, manganese, mouse, remyelination

Introduction

Over the past two decades, magnetic resonance imaging (MRI) has gained tremendous importance – not only in the clinic but also in animal research. This particularly applies to the field of neuroscience and studies of the intact living brain, for example see Natt and Frahm, 2004. Instead of using a human whole body MRI system, most animal studies are better served with a dedicated small-bore magnet and a suitable array of radio frequency antennae for MRI signal excitation and detection.

Apart from being non-invasive, MRI not only offers access to anatomical structures at high spatial resolution using true three-dimensional (3D) image acquisition, but also yields excellent soft-tissue contrast and marked sensitivity to pathological tissue alterations. Specificity is slightly less pronounced, but at least in part compensated for by a variety of different MRI techniques, which complement morphological assessments by giving insight into brain metabolism and function. While applications in animal research range from monkeys to insects, for example see Boretius et al., 2004 and Michaelis et al., 2005, studies of non-human primates or their immediate precursors are rare and only prompted by the need for immunological or behavioural properties that are similar to those of humans. The vast majority of animal MRI studies focuses on rodents.

In general, MRI turns out to be particularly attractive for animal experimentation. First of all, the approach reduces the number of animals required in follow-up studies of disease progression or therapy. Secondly, MRI refines the quality of the data by yielding *in vivo* rather than *post mortem* parameters and by allowing the monitoring of *individual* animals without necessarily relying on group comparisons. Thirdly, studies of mutant animals offer the chance to link molecular biology and neurogenetics to physiological and

functional properties at the system level. And finally, clinical MRI studies may be complemented by new approaches not yet or not directly applicable to humans. The purpose of this work is to illustrate these concepts by a couple of selected applications in mice.

Animals

All studies were performed in accordance with the European Council Directive (86/609/EEC) and approved by the responsible governmental authority (*Bezirksregierung Braunschweig*). Animals were individually housed under conventional conditions in macrolon cages according to the recommendations of the Society for Laboratory Animal Science (Germany).

Magnetic resonance imaging

Mice were anaesthetised by i.p. injection of medetomidine (1.5 mg/kg) and ketamine (150 mg/kg) for induction. Subsequently, animals were intubated with a polyethylene endotracheal tube and artificially ventilated. Anaesthesia was maintained using 0.5-1.0% isoflurane in oxygen and respiration was monitored by a signal derived from a homemade pressure transducer fixed to the animal's chest. For MRI the animals were placed in a prone position on a purpose-built holder with an adjustable nose cone. The rectal temperature was held constant at 37 ± 1°C by covering the body with water blankets. Figure 1 shows the experimental set-up for MRI studies of mouse brain. During the examination, the holder is positioned in the centre of a horizontal magnet. The use of inhalation anaesthesia allows examination for up to 3 hours as well as repeated follow-up studies of individual animals.

All MRI studies were conducted at 2.35 T using an MRBR 4.7/400-mm magnet (Magnex, Scientific, Abington, UK) equipped with a DBX system (Bruker Biospin, Ettlingen, Germany). Radio frequency excitation was accomplished by use of a Helmholtz coil (100 mm) and combined with an elliptical surface coil (20 x 12 mm) for MRI signal detection. Three-dimensional T1-weighted (3D FLASH, repetition time TR = 17 ms, echo time TE = 7.58 ms, flip angle 25°) and T2-weighted images (3D FSE, TR/TE = 3000/98.25 ms, 16 echoes, echo-spacing = 12.5 ms) were obtained with an isotropic spatial resolution of 117 µm (Natt et al., 2002). Maps of the magnetisation transfer ratio (MTR) were based on a spin density-weighted sequence (3D FLASH, TR/TE = 30/7.60 ms, flip angle 5°) with and without off-resonance radio frequency irradiation (frequency offset: 5 kHz, mean amplitude: 200 Hz corresponding to a flip angle of 1045°) (Natt et al., 2003).

Results and discussion

High-resolution 3D MRI of mouse brain *in vivo*

High-quality three-dimensional images of the brain of anaesthetised mice may be obtained at almost 100 µm isotropic spatial resolution within measuring times of 1-2 hours. Figure 2 shows selected sections from T1-weighted 3D MRI acquisitions of two different mouse strains depicting pronounced differences in ventricular volume (Natt et al., 2002). MRI provides reliable access to small structures including the hippocampal formation or the habenular nuclei – truly in three dimensions and with the possibility of arbitrary image reconstructions. Complementary

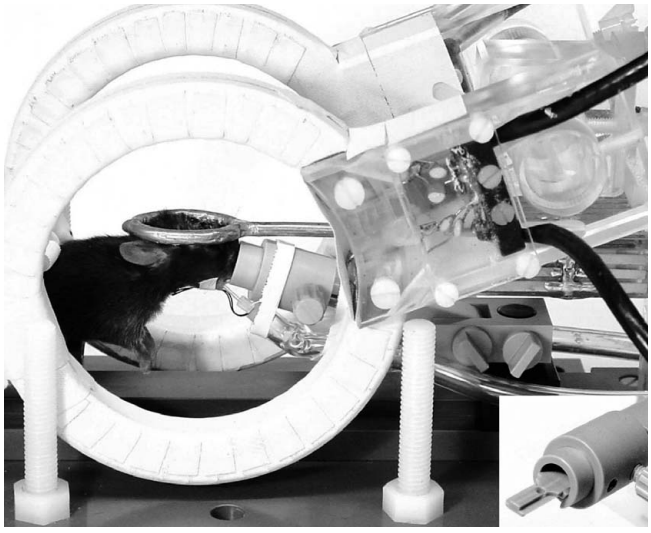


Fig. 1: Experimental set-up for MRI of mouse brain *in vivo*.
Animals are intubated for inhalation anaesthesia and warmed by water-filled blankets. MRI signal excitation and detection is accomplished by use of a large Helmholtz coil (100 mm diameter) surrounding the whole animal and a small elliptical surface coil (20 mm anterior-posterior, 12 mm left-right) atop the head, respectively. The insert (right bottom corner) depicts the head holder comprising a nose cone and bite bar.

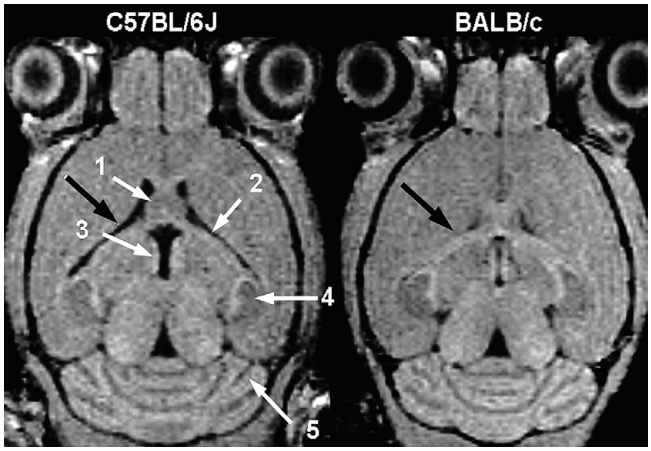


Fig. 2: High-resolution T1-weighted MRI of mouse brain *in vivo*.
The images represent selected sections of a 3D MRI acquisition at 117 µm isotropic resolution (for details see text) in two different strains. C57BL/6J and BALB/c mice present with structural variations such as relatively large ventricles in the forebrain of C57BL/6J animals (black arrows). 1 = Lateral septum, 2 = hippocampal fimbria, 3 = medial habenular nucleus, 4 = hippocampal formation, 5 = cerebellar folia. Adapted from (Natt et al., 2002).

to images with T1 contrast as shown here, many applications rely upon additional T2-weighted images because of their particular sensitivity to detect pathological tissue alterations (see below).

Cuprizone mouse model: MRI of changes in myelination

Mouse models are extensively used to study the pathophysiological mechanisms of multiple sclerosis (MS). This is largely because human MRI studies of MS lesions are hampered by the occurrence of pronounced heterogeneity and the simultaneous presence of inflammation, demyelination (and remyelination), and axonal damage as the pathologic hallmarks of this autoimmune disorder.

Complementary to the widespread use of animal models of experimental allergic encephalitis (EAE), the cuprizone mouse model specifically addresses the aspects of demyelination and remyelination. Feeding mice with cuprizone leads to a selective and reversible demyelination of the corpus callosum with little or no axonal damage (Johnson and Ludwin, 1981; Matsushima and Morell, 2001). The effect has been demonstrated by histological staining for myelin and is also well characterised by MRI as shown in figure 3 (Merkler et al., 2005). In comparison to controls with normal myelination, as for example visualised by a dark corpus callosum on T2-weighted images, demyelination after 6 weeks of cuprizone causes a reversal of contrast, that is a bright intensity on T2-weighted images. After another 6 weeks of normal diet, the animals exhibit partial remyelination of white matter and a respective partial recovery of MRI contrasts. It could be shown that a combination of all three MRI contrasts depicted in figure 3, namely T2, T1, and magnetisation transfer contrast, correctly differentiated 95% of all animals as myelinated, demyelinated, or remyelinated when using a discriminant function analysis (Merkler et al., 2005).

The ability to non-invasively predict the myelin status promises new insights into the evolution of MS lesions. It certainly represents a prerequisite for the *in vivo* monitoring and evaluation of novel therapeutic approaches devised to promote re-myelination. Pertinent strategies include the application of growth factors, the transplantation of myelin-forming stem cells or the use of intravenous immunoglobulin therapies.

Cnp1 deficient mice: MRI of axonal damage

In analogy to the cuprizone mouse model of demyelination, the use of transgenic mice with a *Cnp1* deficiency in oligodendrocytes (Lappe-Siefke et al., 2003) offers selective MRI studies of axonal damage as another pathophysiological mechanism in MS and other neurodegenerative disorders. The gene encodes a protein which turns out to be essential for axonal survival but not for myelin assembly. Accordingly, *Cnp1* deficient mice develop

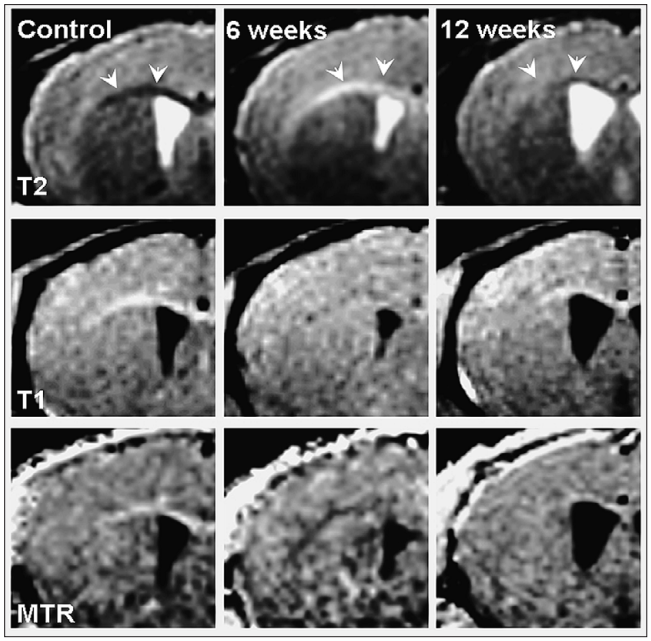


Fig. 3: MRI of myelination in mice treated with cuprizone. In comparison with controls, animals show pronounced demyelination of the corpus callosum (arrowheads) after 6 weeks of cuprizone treatment. The effect is best seen by a reversal of T2 contrast, that shows hypointensities for controls and hyperintensities at 6 weeks, as well as by a reduction of both T1 contrast and the magnetisation transfer ratio (MTR). After an additional 6 weeks on normal diet and withdrawal of the toxin (12 weeks), T1 and T2 contrast as well as MTR partially recover indicating remyelination. Adapted from (Merkler et al., 2005).

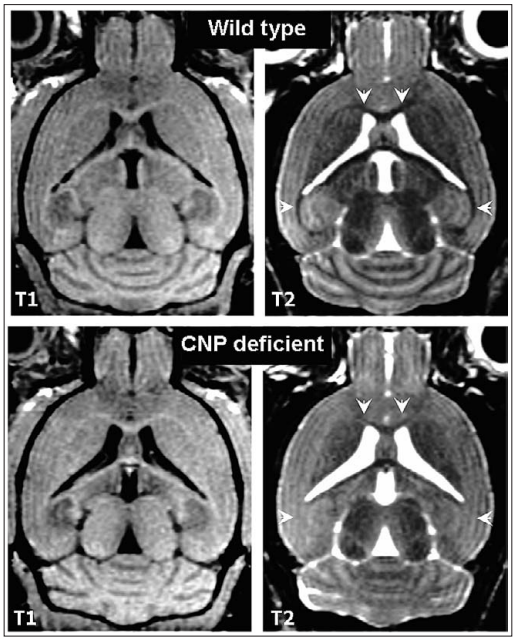


Fig. 4: MRI of axonal damage in *Cnp1*-deficient mice. Mice with a deficiency of the *Cnp1* gene in oligodendrocytes develop axonal damage while maintaining myelin structure and function. *Cnp1* deficient mice at the age of 12 months reveal a reduction or even a complete lack of T2 contrast in the corpus callosum and hippocampal fimbria (arrowheads) without alterations of T1 contrast. This pattern differs from the MRI findings for demyelination (see figure 3).

axonal swelling and neurodegeneration throughout the brain without demyelination.

MRI of *Cnp1* deficient mice reveals no changes in T1-weighted images of white matter structures but a lack of contrast in T2-weighted images. These observations represent a pattern that clearly differs from the MRI findings for demyelination (see previous section). As demonstrated in figure 4, the exclusive T2 changes not only apply to the corpus callosum but also to the hippocampal fimbria. Moreover, the mutants exhibit altered diffusion properties and reduced brain volumes. Together, these preliminary findings may lead to new in vivo markers of axonal damage.

Functional staining of neuroaxonal connectivity using Mn^{2+}-enhanced MRI

Extending structural characterisations, contrast-enhanced MRI aims at functional assessment of synaptic activity and activity-dependent axonal transport in various brain systems. The concept is based on the application of divalent manganese ions (Mn^{2+}) that, similar to calcium ions (Ca^{2+}), are taken up by voltage-gated calcium channels into active neurons. At the same time, Mn^{2+} ions serve as MRI contrast agent, shortening the T1 relaxation time (Lin and Koretsky, 1997; Pautler et al., 1998). Thus, Mn^{2+}-enhanced images with hyperintensities in areas of specific uptake or accumulation may be recorded at high spatial resolution using T1-weighted 3D MRI. As an early example, the intravitreal application of $MnCl_2$ into one eye of a rat delineated the entire visual pathway from the optic nerve, chiasma, contralateral optic tract, and lateral geniculate to the superior colliculus and suprachiasmatic nucleus (Watanabe et al., 2001).

Subsequently, Mn^{2+}-enhanced 'staining' of specific brain systems and their axonal connections has been exploited for functional mapping of hippocampal projections after bilateral intraparenchymal injection (Watanabe et al., 2004a). Figure 5 reveals a regionally selective enhancement within the hippocampal formation that almost exclusively affects the dentate gyrus and CA3 subfield with only very mild extensions into CA1 and the subiculum. This pattern is in line with differences in the excitability of hippocampal neurons previously seen in electrophysiologic recordings. Uptake is mainly in active pyramidal cells, which is further supported by the axonal projections toward the lateral septum shown in figure 6. Taken from the same 3D MRI data set as the image shown in figure 5, sections anterior to the hippocampus delineate the efferent hippocampal pathways to the ventral hippocampal commissure and the septal region by enhancement in the fimbria, ventral hippocampal commissure, precommisural fornix, and septal nuclei (Watanabe et al., 2004a). In conjunction with suitable cognitive tasks and behavioural assessments, high-resolution Mn^{2+}-enhanced MRI of normal and mutant mice is expected to become a powerful tool for a further characterisation of learning and memory processing in the hippocampal system.

The functional relevance of manganese contrast is further evidenced by observations after injection into the lateral ventricle or after systemic subcutaneous administration, in which case Mn^{2+}-uptake into the central nervous system is accomplished via the CSF-brain barrier (Watanabe et al., 2002). Figure 7 shows the uptake of manganese into active neurons of the habenulo-interpeduncular system where the initial enhancement of the habenula is followed by MRI signal increases of the fasciculus retroflexus as the connecting axonal tract and finally the interpeduncular nucleus (Watanabe et al., 2004b). Though not directly visible in a single 3D MRI scan, the sequential enhancement may be demonstrated by repetitive 3D MRI recordings.

While the hippocampal formation is involved in learning and memory, the habenular complex is a crucial relay structure in the diencephalic conduction system and is as such involved in cog-

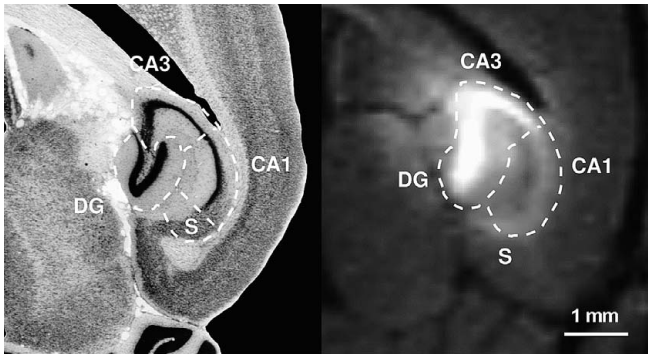


Fig. 5: Manganese-enhanced MRI of the mouse hippocampus.
The mouse hippocampal formation in (left) a histologic section (Cresyl violet, adapted from Rosen et al., 2000) and (right) a corresponding T1-weighted section from a 3D MRI data set 6 hours after intraparenchymal injection of $MnCl_2$ (0.25 µl, 200 mM). Pronounced signal enhancements in the dentate gyrus (DG) and CA3 subfield are complemented by only very mild signal increases in the CA1 subfield and subiculum (S). Adapted from (Watanabe et al., 2004a).

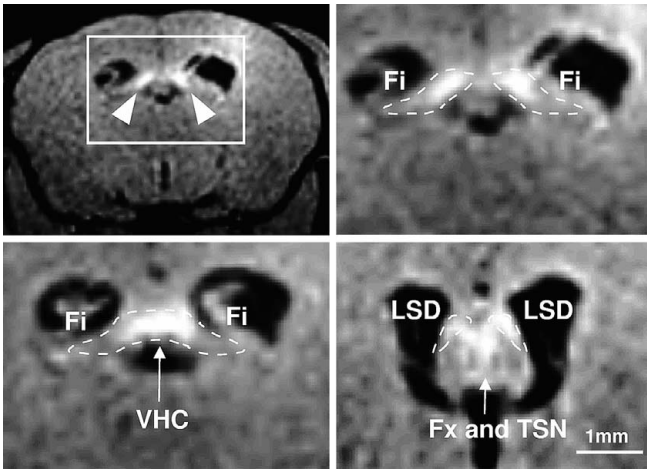


Fig. 6: Manganese-enhanced MRI of the mouse hippocampal projections.
Selected sections (magnified views) along the hippocampo-septal projection pathway 6 hours after intraparenchymal injection of $MnCl_2$ (same T1-weighted MRI data set as shown in figure 5) depict enhanced structures such as the medial part of the fimbria (Fi) bilaterally, the ventral hippocampal commissure (VHC), the dorsal part of the lateral septal nucleus (LSD) bilaterally, the precommissural fornix (Fx), and the triangular septal nucleus (TSN). Adapted from (Watanabe et al., 2004a).

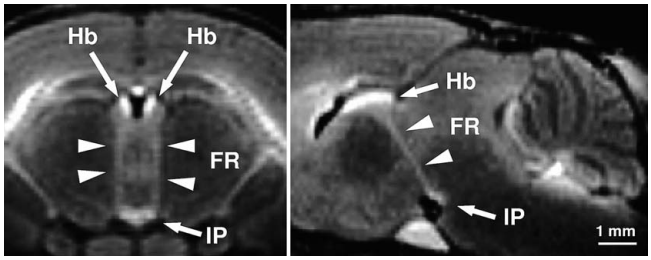


Fig. 7: Manganese-enhanced MRI of the mouse habenulo-interpeduncular system.
The images represent (left) an oblique section and (right) a parasagittal section of a 3D MRI data set 6 hours after subcutaneous injection of $MnCl_2$ (5 ml/kg, 120 mM). They demonstrate enhancement in the anterior parts of the habenulae (Hb), the fasciculus retroflexus (FR) as its major axonal connection, and the interpeduncular nucleus (IP). Adapted from (Watanabe et al., 2004b).

nition and behaviour, for example as part of the reward system. As an example, functional Mn^{2+}-enhanced MRI may therefore offer completely new ways to study mutant mice with – or without – recently discovered susceptibility genes for neuropsychiatric diseases such as schizophrenia.

Conclusion

The most prominent advantage of MRI is the fact that the procedure itself is non-invasive. MRI helps to reduce the number of animals used and improves the quality of the research by yielding *in vivo* rather than *post mortem* results, as well as by monitoring disease progression or therapeutic efficacy in *individual* animals. Multiple techniques offer insights into the anatomy, metabolism, and function of the central nervous system for a wide range of species. Thus, MRI contributes to bridging basic and clinical neuroscience and emerges as a key tool in translational research.

References

Boretius, S., Natt, T., Watanabe, T., Ehrenreich, L., Tammer, R., Frahm, J. and Michaelis, T. (2004). In vivo diffusion tensor mapping of the brain of squirrel monkey, rat, and mouse using single-shot STEAM MRI. *MAGMA 3*, 1-9.

Johnson, E.S. and Ludwin, S.K. (1981) The demonstration of recurrent demyelination and remyelination of axons in the central nervous system. *Acta Neuropathol. 53*, 93-98.

Lappe-Siefke, C., Goebbels, S., Gravel, M., Nicksch, E., Lee, J., Braun, P.E., Griffiths, I.R. and Nave, K.A. (2003) Disruption of Cnp1 uncouples oligodendroglial functions in axonal support and myelination. *Nat. Gen. 33*, 366-374.

Lin, Y.J. and Koretsky, A.P. (1997) Manganese ion enhances T_1-weighted MRI during brain activation: an approach to direct imaging of brain function. *Magn. Reson. Med. 38*, 378-388.

Matsushima, G.K. and Morell, P. (2001) The neurotoxicant, cuprizone, as a model to study demyelination and remyelination in the central nervous system. *Brain. Pathol. 11*, 107-116.

Merkler, D., Boretius, S., Stadelmann, C., Ernsting, T., Michaelis, T., Frahm, J. and Brück, W. (2005). Multicontrast MRI of remyelination in the central nervous system. *NMR Biomed. 18*, 395-403.

Michaelis, T., Watanabe, T., Natt, O., Boretius, S., Frahm, J., Utz, S. and Schachtner, J. (2005). In vivo 3D MRI of insect brain: Cerebral development during metamorphosis of Manduca Sexta. *NeuroImage 24*, 596-602.

Natt, O., Watanabe, T., Boretius, S., Radulovic, J., Frahm, J. and Michaelis, T. (2002). High-resolution 3D MRI of mouse brain reveals small cerebral structures in vivo. *J. Neurosci. Meth. 120*, 203-209.

Natt, O., Watanabe, T., Boretius, S., Frahm, J. and Michaelis, T. (2003). Magnetization transfer MRI of mouse brain reveals areas of high neural density. *Magn. Reson. Imaging 21*, 1113-1120.

Natt, O. and Frahm, J. (2004). In vivo magnetic resonance imaging: Insights into structure and function of the central nervous system. *Meas. Sci. Technol. 16*, R17-R36.

Pautler, R.G., Silva, A.C. and Koretsky, A.P. (1998) In vivo neuronal tract tracing using manganese-enhanced magnetic resonance imaging. *Magn. Reson. Med. 40*, 740-748.

Rosen, G.D., Williams, A.G., Capra, J.A., Connoly, M.T., Cruz, B., Lu, L., Airey, D.C., Kulkarni, K. and Williams, R.W. (2000). The mouse brain library@www.mbl.org Int. Mouse *Genome Conf. 14*, 166.

Watanabe, T., Michaelis, T. and Frahm, J. (2001). Mapping of retinal projections in the living rat using high-resolution 3D gradient-echo MRI with Mn^{2+}-induced contrast. *Magn. Reson. Med. 46*, 424-429.

Watanabe, T., Natt, O., Boretius, S., Frahm, J. and Michaelis, T. (2002). In vivo 3D MRI staining of mouse brain after subcutaneous application of $MnCl_2$. *Magn. Reson. Med. 48*, 852-859.

Watanabe, T., Radulovic, J., Spiess, J., Natt, O., Boretius, S., Frahm, J. and Michaelis, T. (2004a). In vivo 3D MRI staining of the mouse hippocampal system using intracerebral injection of $MnCl_2$. *NeuroImage 22*, 860-867.

Watanabe, T., Frahm, J. and Michaelis, T. (2004b). Functional mapping of neural pathways in rodent brain in vivo using manganese-enhanced three-dimensional magnetic resonance imaging. *NMR Biomed. 17*, 554-568.

Correspondence to

Prof. Dr. Jens Frahm
Biomedizinische NMR Forschungs GmbH
Max-Planck-Institut für biophysikalische Chemie
Am Fassberg 11
37070 Göttingen, Germany
e-mail: jfrahm@gwdg.de

Session 2.4
Non-human primates – housing, enrichment, positive reinforcement training

Primates in Laboratories: Standardisation, Harmonisation, Variation and Science

Hannah M. Buchanan-Smith
Department of Psychology, University of Stirling, Scotland, UK

Summary
Many animals are bred specifically for use in laboratories; the genetic variation between individuals is minimised, and housing and husbandry is often standardised. The rationale is to decrease the variation of the scientific findings, and allow a reduction in number of animals used, although these relationships are questioned. Non-human primates used in laboratories present a different case; there are genetic differences, and considerable variation in rearing practices, housing, enrichment and training both among, and often within, facilities. The impact of this variation on both welfare and science warrants further investigation.

Keywords: three Rs, welfare, enrichment

Introduction

Animals are housed in laboratories for scientific research and testing. Russell and Burch's three Rs (replacement, reduction and refinement) are now widely accepted, and provide a systematic framework to progress humane animal use. Setting standards, standardisation and harmonisation have the potential to impact on science and on the three Rs. The aims of this paper are to:

- clarify the terms "setting standards", "standardisation" and "harmonisation", and to summarise their potential impact on the three Rs and science;
- describe how housing and husbandry practices vary for non-human primates (hereinafter primates);
- argue that the impact of such variation should be quantified with the goal of striving towards well-adjusted animals for sound scientific results.

Agreeing upon definitions of terms

There is a need to agree upon definitions before one can proceed! In the 45 years since Russell and Burch (1959) first proposed the three Rs, there have been considerable advances in animal ethics and animal welfare science. Despite the progressive nature of the three Rs as described by Russell and Burch, a recent analysis found over 15 different definitions of refinement, some of which are regressive from the intended meaning. A suggested definition has been proposed: "*any approach which avoids or minimises the actual or potential pain, distress and other adverse effects experienced at any time during the life of the animals involved, and which enhances their wellbeing*" (Buchanan-Smith et al., 2005). This not only makes it explicit that refinement applies to all stages of the life of an animal destined for use in the laboratory (i.e. the birth to death experience), but also takes a proactive stance of enhancing wellbeing. There has been less confusion about the meaning of replacement and reduction and Russell and Burch's original definitions of "*any scientific method employing non-sentient material which may, in the history of experimentation, replace methods which use conscious living vertebrates*" and "*reduction in the number of animals used to obtain information of a given amount and precision*" (Russell and Burch, 1992) are used here.

A further confusion in the literature is that standardisation has been interpreted in three ways. The first is the setting of stan-

dards – referred to here as "setting standards". The second is defining the properties of experimental animals and their environments to increase the reproducibility of results (Beynen et al., 2001), referred to here as "listing" (see Würbel, 2002). The third refers to keeping experimental conditions the same for all animals (within-experimental standardisation) or all experiments (between-experiment standardisation – see Würbel, 2002). In this paper, the latter will be referred to as standardisation. The difference between setting standards and harmonisation can also be confused (see below).

Setting standards and harmonisation

There is debate about the need for setting standards and harmonisation in laboratory animal science (e.g. ILAR, 2003). One of the problems is they can refer to a multitude of levels (e.g. opinions, animal welfare, environments, practices, processes, scientific protocols, tests), and whether they are a desirable goal surely lies with the proposed level, taking into account the broader context, including the ability to implement, and the means of ensuring the desired outcome has been fulfilled.

One interpretation of the difference between setting standards and harmonisation is that harmonisation refers to performance outcomes whilst setting standards is prescriptive and refers specifically to how to get to the outcome. So for example, a legal requirement to adhere to minimum cages sizes based on body weight of the animal is a standard, and standards of this sort are set in an attempt to ensure a minimum level of welfare. They can be easily quantified, and monitored, which constitutes an advantage for legal oversight/regulation. Such prescriptive standards should be based, wherever possible, on scientific findings or at least well-established practice. It is also important to note that they should be treated as the minimum required and providing more than these minimum standards is encouraged to enhance welfare (i.e. minimum standards do not necessarily imply standardisation).

In contrast, harmonisation of acceptable welfare might refer to providing animals with cage dimensions that allow them to meet species-specific and individual needs. For example, Buchanan-Smith et al. (2004) have argued that setting minimum cage sizes for primates based on body weight alone is inappropriate to ensure acceptable welfare, but that a range of characteristics such as morphometric, physiological, ecological, locomotor, social, reproductive and behavioural characteristics, together with the primate's age, sex and individual history should be taken into account. Harmonisation is therefore less prescriptive than minimum standards and allows for greater flexibility in how to achieve the goal of enhancing welfare.

This flexibility can be beneficial to allow innovation and to challenge and improve conditions. Further it encourages one to take the individual needs of animals into account. This flexibility may contribute positively to refinement but flexibility may be inappropriate due to discrepancy in the interpretation of acceptable welfare as a consequence of differences among cultures, religions, legislation, and regulations across countries. Therefore specific guidance, and specific standards, may need to be imposed and strictly monitored to increase the likelihood that there is a minimum standard of animal care that is adhered to on an international basis.

Protocols and legally required safety tests may also be harmonised internationally. This not only represents an important strategy to reduce the number of animals used in experimentation in individual countries, but it also helps to ensure that obsolete invasive animal studies are replaced by the most advanced techniques using insentient material (de Boo et al., 2005). Therefore there is the potential to act positively on both reduction and replacement. Further, setting standards and harmonising animal care and use internationally has been viewed as a socio-political and economic imperative, essential to the elimination of trade barriers and to promote multinational interactions (Miller, 1998).

As already mentioned, there are two further interpretations of "standardisation". The benefits of "listing" and "standardisation" are also debated (see Würbel, 2002). For example, rigorous attempts to equate test apparatus, testing protocols, and all possible features of animal husbandry have not succeeded. Crabbe et al. (1999) found that there were significant differences between the behaviour of mice of identical strains, across three test laboratories, and their findings demonstrate the difficulty, or indeed impossibility, of standardisation to guarantee reproducibility of results across experiments. Würbel (2002) argues that standardisation is done at the expense of external validity. By limiting the experiment to the effects of the variable under question on a sample of, for example, a specific genetic strain, under specific standardised conditions, the generalisability of the results to other conditions, populations or species is limited.

Variation

From the discussion above, we get a flavour of the debate and the potential impact of setting standards, standardisation and harmonisation on the three Rs, and on science. However, there is a complex interplay between the three Rs (de Boo et al., 2005), and there is also the potential for standardisation to create a conflict between reduction and refinement. For example, providing housing to enhance individual wellbeing will have a positive impact on refinement, but the potential for a negative impact on reduction. One reason is because enriched housing may confound experimental results (enrichment is known to impact on a wide range of parameters), increasing variance and the number of animals required (e.g. Bayne, 2005; Benefiel et al., 2005; Mering et al., 2001). A negative consequence of such reasoning is that some scientists are reluctant to provide enrichment for their research animals. However, although there are some studies that have concluded that enrichment increases variation in results (e.g. see review in Van de Weerd et al., 2002) a most systematic and comprehensive study has shown that enrichment does not disrupt standardisation and reproducibility of results (Wolfer et al., 2004). Wolfer et al. (2004) compared female mice housed in barren cages with those in large enriched cages on a range of behavioural tests. They replicated their tests three times, in each of three laboratories. They conclude that housing

in enriched cages, which should improve the wellbeing, does not increase variation in the findings, and should therefore be used as the norm. This conclusion is supported by Garner (2005) who argues that abnormal repetitive behaviours, which are seen more often in barren conditions, is likely to compromise validity and reproducibility of results.

Primates in laboratories

Whilst the debate about whether enrichment is a desirable goal for non-primates (e.g. rodents, rabbits etc.) used in laboratories continues, due to the potential effects on science (e.g. Benefiel et al., 2005), primates used in laboratories often present a very different case. The most commonly used primates are the rhesus macaque (*Macaca mulatta*), the long-tailed or cynomolgus macaque (*M. fascicularis*) and the common marmoset (*Callithrix jacchus*) (see Rennie and Buchanan-Smith (2005) for an analysis of primate use in Europe). Primate use makes up a very small percentage of total animal use in scientific research and testing. However, due to the genetic similarity with humans, their intelligence, sentience and sociality, the complexity of their physical as well as social needs, and the clear temperament differences both among (e.g. Clarke et al., 1988) and within species (e.g. Capitanio, 2004) they have been treated differently to other animals used in laboratories.

Sources of variation for primates

Despite the potential benefits of setting standards, standardisation and harmonisation described above, for primates there is great variation in, for example, sourcing, rearing practices, weaning policy, housing, enrichment and training both among, and often within, facilities. Below these differences are briefly outlined, and it is highlighted that if the harmonised goal is to produce well adjusted primates suitable for good scientific research, the impact of this variation must be investigated.

Although common marmosets are generally bred in-house, macaques are sourced from many different countries, and often bred in countries where the macaques occur naturally (for UK supply see Prescott, 2001). There are several, potentially important differences between individuals from these overseas establishments, and those that are bred in the user establishment. For example, those bred overseas are often first generation from wild- caught macaques, whilst those bred in house are from captive-bred parents. Those from overseas are likely to have had greater exposure to diseases, and there is less potential for preparation for future use (e.g., socialisation to humans, positive reinforcement training). Further, if bred in house, there may be no, or limited, transport required, whilst those from overseas often have long transport and relocation, which is known to impact negatively on welfare (Honess et al., 2004). Furthermore there are known genetic differences between macaques from different habitats (Zhang and Shi, 1993), yet, unlike the reporting of the exact strain of the rodent used, few details are given in scientific articles on the source of the primate used in the experiment.

Similarly, there is considerable variation in weaning policy, when offspring are removed from their mothers. Most guidelines refer to age as the sole criterion and the guidelines of minimum weaning age vary considerably (6-12 months, Prescott et al., in prep.) as do policies between overseas establishments (e.g. from 90 days, Welshman, 1999). It is well known that early weaning has serious consequences for behaviour, physiology and immune function. Again, like body weight as a measure for determining minimum cage sizes, age has the advantage of being unambiguously quantified, but in order to achieve the harmonised goal of socially well adjusted primates that can cope with the challenges they will face in laboratory experimentation, a range of measures should be used. These might include age, in addition to weight, behaviour and physiological measures, and any guidance should be based on sound evidence.

Given the known impact of early separation, further research is required on the use of rotational hand-rearing, a practice that is used widely in marmosets. Although twins are the norm in the wild, triplet births are increasingly common in captivity, probably due to the rich diet (Tardif and Jaquish, 1997). In some laboratories, the triplets are rotationally hand-reared. In these instances, two infants are left with the marmoset family, whilst one is removed for hand-feeding on a rotational basis. Occasionally, mothers will die or reject their infants and the offspring are entirely human hand-reared. Even brief early social deprivation (9 hours/week on postnatal days 2-28) is known to adversely affect endocrine and behavioural responses to psychosocial challenges in the long-term (Dettling et al., 2002a, 2002b). This suggests that the highly productive management practice of rotational hand-rearing which necessitates early deprivation from the family for long periods is likely to have a significant impact on experimental results and should be quantified.

Housing, husbandry and enrichment may also impact on experimental results. There is a vast literature on the diversity of "enrichment" that has been provided for primates in laboratories. These include housing (e.g. the provision of outside runs and exercise cages), structural (e.g. perches and swings, "toys" and manipulanda, visual barriers, water baths), feeding (e.g. grooming boards, puzzle feeders, gum trees for marmosets), visual and auditory (e.g. television or video, computer tasks, music, mirrors), and social enrichment (e.g. pair or group housing) (e.g. see Honess and Marin, 2006; Lutz and Novak, 2005). However, although the benefits of such enrichment in enhancing welfare are regularly reported, no comprehensive attempt has been made to determine how different enrichment techniques impact upon variation in experimental findings. In this respect, the primate literature lags behind that for other animals used in laboratories.

A final example of variation is from the increasing use of positive reinforcement to train primates to co-operate during laboratory procedures. Although there is little doubt that this practice is beneficial (e.g. see Prescott and Buchanan-Smith, 2003) in terms of reduction of stress, improvement in welfare and facilitation of experimental protocols, there is considerable variation in its uptake (Prescott et al., 2005). Further, standardisation of overall length of time that animals are trained for may be

inappropriate as primates learn at different rates, and temperament differences will impact on training success (e.g. Coleman et al., 2005). As Videan et al. (2005) demonstrate, chimpanzees that have been trained over 21 months vary considerably in what they can be trained to do, and in the time taken to be successfully trained. Not all are able to be trained for certain tasks, and those successfully trained to present for injection do not necessarily have lower stress levels than those that could not be trained and requiring darting. The level of co-operation was reported to be a more important factor in determining the amount of stress experienced during anaesthetisation (Videan et al., 2005). Again the impact on science is not well understood, but a desirable goal would be harmonisation of training procedures to ensure a similar level of co-operation is achieved. Similar principles apply to habituation to novel environments or tests procedures (such as restraint chairs).

Conclusions

Primates have, to some extent, escaped the same genetic, rearing and environmental standardisation that is seen with the majority of laboratory animals. This has been a mixed blessing. On the one hand, there has been far greater innovation in enrichment (e.g. outside runs, manipulanda, visual and auditory enrichment) designed to enhance wellbeing, and these have been implemented liberally. On the other hand, this variation in housing, together with the numerous different rearing conditions, weaning policies, husbandry and other practices have not been systematically quantified and their impact on the scientific outcome is not known. They may increase variation in results, and thus more primates may be used in experiments than are needed (conflicting with reduction), although the external validity (generalisability) of the findings might be increased. The impact of such variation certainly warrants systematic investigation. It is concluded that greater harmonisation is needed in primate use and care, with the goal of ensuring socially well adjusted primates that are able to cope with the challenges they will face in experimentation and to ensure sound scientific results.

References

Bayne, K. (2005). Potential for Unintended Consequences of Environmental Enrichment for Laboratory Animals and Research Results. *ILAR Journal 46*, 129-139.

Benefiel, A. C., Dong, W. K. and Greenough, W. T. (2005). Mandatory "Enriched" Housing of Laboratory Animals: The Need for Evidence-based Evaluation. *ILAR Journal 46*, 95-105.

Beynen, A. C., Gartner, K. and Van Zutphen, L. F. M. (2001). Standardization of animal experimentation. In L. F. M. Van Zutphen, V. Baumans and A. C. Beynen (eds.), *Principles of Laboratory Animal Science, revised edition* (103-110). Amsterdam: Elsevier.

Buchanan-Smith, H. M., Rennie, A. E., Vitale, A. et al. (2005). Harmonising the Definition of Refinement. *Animal Welfare 14*, 379-384.

Buchanan-Smith, H. M., Prescott, M. J. and Cross, N. J . (2004). What factors should determine cage sizes for primates in the laboratory? *Animal Welfare 13 (Suppl.)*, 197-201.

Capitanio, J. P. (2004). Personality factors between and within species. In B. Thierry, M. Singh and W. Kaumanns (eds.), *Macaque Societies: A Model for the Study of Social Organization* (13-33). New York: Cambridge University Press.

Clarke, A. S., Mason, W. A. and Moberg, G. P. (1988). Differential behavioral and adrenocortical responses to stress among three macaque species. *American Journal of Primatology 14*, 37-52.

Coleman, K., Tully, L. A. and McMillan, J. L. (2005). Temperament correlates with training success in adult rhesus macaques. *American Journal of Primatology 65*, 63-71.

Crabbe, J. C., Wahlsten, D. and Dudek, B. C. (1999). Genetics of mouse behavior: Interactions with laboratory environment. *Science 284*, 1670-1672.

de Boo, M. J., Rennie, A. E., Buchanan-Smith, H. M. and Hendriksen, C. F. M. (2005). The Interplay between Replacement, Reduction and Refinement. *Animal Welfare 14*, 327-332.

Dettling, A. C., Feldon, J. and Pryce, C. R. (2002a). Early deprivation and behavioral and physiological responses to social separation/novelty in the marmoset. *Pharmacology, Biochemistry and Behavior, 73*, 259-269.

Dettling, A. C., Feldon, J. and Pryce, C. R. (2002b). Repeated parental deprivation in the infant common marmoset (*Callithrix jacchus*, Primates) and analysis of its effects on early development. *Biological Psychiatry 52*, 1037-1046.

Garner, J. P. (2005). Stereotypies and Other Abnormal Repetitive Behaviors: Potential Impact on Validity, Reliability, and Replicability of Scientific Outcomes. *ILAR Journal 46*, 106-117.

Honess, P. E., Johnson, P. J. and Wolfensohn, S. E. (2004). A study of behavioural responses of non-human primates to air transport and re-housing. *Laboratory Animals 38*, 119-132.

Honess, P. E. and Marin, C. M. (2006). Enrichment and aggression in primates. *Neuroscience and Biobehavioral Reviews 30,* 413-436.

Institute for Laboratory Animal Research (2004). *The Development of Science-Based Guidelines for Laboratory Animal Care*. Proceedings of the November 2003 International Workshop, The National Academies Press, Washington, D.C.

Lutz, C. K. and Novak, M. A. (2005). Environmental Enrichment for Nonhuman Primates: Theory and Application. *ILAR Journal 46*, 178-191.

Mering, S., Kaliste-Korhonen, E. and Nevalainen, T. (2001). Estimates of appropriate number of rats: Interaction with housing environment. *Lab Animal 35*, 80-90.

Miller, J. (1998). International harmonization of animal care and use: The proof is in the practice. *Lab. Animal 27*, 28-31.

Prescott, M. J. (2001). Counting the cost: *Welfare implications of the acquisition and transport of non-humans primates for use in research and testing*. Horsham, U.K.: RSPCA.

Prescott, M. J. and Buchanan-Smith, H. M. (eds.) (2003). Training nonhuman primates using positive reinforcement

techniques. *Journal of Applied Animal Welfare Science 6*, 157-261.

Prescott, M. J., Rennie, A. E. and Buchanan-Smith, H. M. (2005). Training of laboratory-housed non-human primates in the UK. *Anthrozoos 18,* 288-303.

Prescott, M. J., Bard, K., Brend, S. et al. (in prep.). An investigation of the scientific basis for weaning policy for macaques bred for use in research and testing.

Rennie, A. E. and Buchanan-Smith, H. M. (2005). Report on the extent and character of primate use in scientific procedures across Europe in 2001. *Laboratory Primate Newsletter 44*, 6-12.

Russell, W. M. S. and Burch, R. L. (1959). *The principles of Humane Experimental Technique*. London: Methuen.

Russell, W. M. S and Burch, R. L. (1992). *The Principles of Humane Experimental Technique*. Universities Federation for Animal Welfare (2nd printing). Herts, England.

Tardif, S. D. and Jaquish, C. E. (1997). Number of ovulations in the marmoset monkey (*Callithrix jacchus*): Relation to body weight, age and repeatability. *American Journal of Primatology 42*, 323-329.

Van de Weerd, H. A., Aarsen, E. L., Mulder, A. et al. (2002). Effects of environmental enrichment for mice: Variation in experimental results. *Journal of Applied Animal Welfare Science 5*, 87-109.

Videan, E. N., Fritz, J., Murphy, J. et al. (2005). Does training chimpanzees to present for injection lead to reduced stress? *Laboratory Primate Newsletter 44*, 1-2.

Welshman, M. (1999). Breeding macaques in source countries. In T. Poole and P. English (eds.), *The UFAW Handbook on the Care and Management of Laboratory Animals*, 7th Edition, Vol 1: Terrestrial Vertebrates (636-642), Oxford: Blackwell Science.

Wolfer, D. P., Litvin, O., Morf, S. et al. (2004). Cage enrichment and mouse behaviour: test responses by laboratory mice are unperturbed by more entertaining housing. *Nature 432*, 821-822.

Würbel, H. (2002). Behavioral phenotyping enhanced – beyond (environmental) standardization. *Genes, Brain and Behavior 1*, 3-8.

Zhang, Y. P. and Shi, L. M. (1993). Phylogeny of rhesus monkeys (*Macaca mulatta*) as revealed by mitochondrial DNA restriction enzyme analysis. *International Journal of Primatology 14*, 587-605.

Acknowledgements

I thank Drs Mark Prescott and Hanno Würbel for helpful comments on the manuscript.

Correspondence to

Dr. Hannah M. Buchanan-Smith
Scottish Primate Research Group
Department of Psychology
University of Stirling
Stirling FK9 4LA
Scotland UK
e-mail: h.m.buchanan-smith@stir.ac.uk

Theme 3
Moral issues of animals, alternatives and public policy

Session 3.1
Influencing and making public policy

Funding for Research, Development, Validation and Acceptance of Alternatives Must Become Transparent

Sara J. Amundson
Doris Day Animal League, Washington, DC, USA

Summary
Attempts to ascertain the amount of funding available worldwide for research, development, validation and, ultimately, acceptance of non-animal and other alternatives have been stunted by the lack of coordinated accounting and a centralised database. In order to prioritise replacement methods and also identify immediate needs in the field of alternatives, an electronic document or database must be collated on an annual basis to provide genuine transparency of federal and other funding. Federal regulatory and research agencies in the United States, European Union, Japan and other countries should agree to a pilot project. Additional funding data should be provided by the regulated industries, public and private scientific organisations and the animal advocacy community.

Keywords: funding, non-animal methods, alternative methods

Introduction

The purpose of this paper is to provide a window into the funding of research, development and validation of non-animal and other alternatives by some federal governments, the regulated industry, independent scientific institutions and the animal advocacy community. The relative paucity of information reveals the necessity for a coordinated, annual approach to collating and categorising accurate information. Accurate information regarding the amount of funding, the projects funded and their direct application to replacing, refining or reducing animals in regulatory toxicity testing is necessary to provide a mechanism for measuring commitment and success in this area.

Methods

The primary methods for obtaining information were websites and direct communications with federal regulators, members of industry and animal advocacy organisations. While industry, the private and public science sector and animal advocacy organisations tend to track annual funding, it is difficult, if not currently impossible, to break out the spending by federal agencies in the United States. The European Union provides more specific information, but still does not readily itemise the particular projects. Contacts with Japan did not yield definitive results.

Results

The Federal Governments – Regulatory and Research Agencies: United States Environmental Protection Agency – The United States Environmental Protection Agency (EPA) serves as one case study in this area. After protracted negotiations with the animal advocacy community, led by the Doris Day Animal League and People for the Ethical Treatment of Animals, the EPA provided $500,000 over a two-year period for research, development and validation of non-animal, alternative test methods. In addition, the National Institute of Environmental Health Sciences committed to $4.5 million over the same fiscal years, 2000-2001.[1]

Following this commitment, the Doris Day Animal League and other animal protection organisations lobbied the United States Congress to secure an appropriation of $4 million in the EPA's Science and Technology Account under the Office of Research and Development for fiscal year 2003 for research, development and validation of non-animal, alternative test methods. This effort was led by Chairman James Walsh of the House Subcommittee on VA, HUD and Independent Agencies Appropriations and resulted in the first-ever congressional directive to the agency in this area.[2] It should be noted that between the fiscal years of 2003-2006, the EPA's Science and Technology Account has hovered between $500-700 million. This is clearly noteworthy as the fiscal year commitment of $4 million was less than one percent of the total research budget.

The EPA's Computational Toxicology Program received strong support from Congress through increased appropriations until fiscal year 2006. Certainly, it can be contended that a portion of this budget under the Office of Research and Development at the agency is directed at non-animal and other alternative test method development and potentially validation.[3]

Unfortunately, even after numerous attempts by animal protection organisations and letters from members of Congress requesting the EPA provide an accounting of how the various funds were expended, a satisfactory explanation has not been received. However, report language in the fiscal year 2006 EPA appropriations bill supported by the American Chemistry Council and the Doris Day Animal League specifically requires the agency to report back to Congress by early 2006 with an accounting of these expenditures.

United States National Institute of Environmental Health Sciences (NIEHS) – The entire fiscal year 2005 budget for NIEHS, including the National Toxicology Program (NTP), hovers at approximately $650 million, which is a small portion of the overall EPA budget. In fact, it is merely equal to EPA's research budget.[4] The significance of this disparity in funding is that the NTP is relied on by most United States federal regulatory and research agencies as the entity to characterise both acute and chronic exposures to chemicals, including a vast carcinogenicity database.

The NIEHS houses the National Interagency Center for Alternative Methods (NICEATM) for the Interagency Coordinating Center for the Validation of Alternative Methods (ICCVAM), which was permanently codified by Public Law No. 106-545.[5] Essentially, the ICCVAM receives its funding from two sources: NIEHS provides approximately $2.6 million per fiscal year; and federal agencies may contract for work through interagency agreements, much like EPA did with *in vitro* methods for its Endocrine Disruptor Screening Program. As the ICCVAM's function is not research and development, but evaluation of alternative and new or revised animal tests, its needs will arise with the number of methods in its pipeline.

European Union – The European Center for the Validation of Alternative Methods (ECVAM) has a broad mandate under existing law. It performs a centralised prioritisation, research, development, validation and recommendations regarding acceptance function. The 2003-2006 European Union Framework Programme 6 funds ECVAM at 35.2 million Euros for validation activities and 22.6 million Euros for methods development.[6] ECVAM is responsible for cultivating and finalising alternative methods for a variety of legislative and regulatory mandates, including the cosmetics directive and the pending REACH programme.

It should be noted that the European Parliament in a November 2004 analysis of the 6th Framework Programme funding for the entity stated, "The political demand for alternative methods is high but must be underlined that ECVAM is only one link in the chain leading from the development of alternative methods to their regulatory acceptance."[7]

To that end, the European Union has a variety of national government centers that focus on alternatives. A few include the Centre for Documentation and Evaluation of Alternatives to Animal Experiments (ZEBET) under the German Bureau for Risk Assessment. ZEBET has expended a total of 5.5 million Euros for research in the alternatives area.[8] Also of note is the United Kingdom's new National Centre for the Replacement, Refinement and Reduction of Animals in Research. It awarded 8 projects for approximately 1 million pounds in 2005.[9]

The regulated industry

Industry has long played a role in funding alternatives development and several of the accepted alternatives had their inception in industry-funded laboratories. A few companies specifically tout their long-term leadership in this area on their websites and in other communications, including The Procter & Gamble Company, Colgate-Palmolive and the United States trade association the Cosmetic, Toiletry and Fragrance Association (CTFA). The Procter & Gamble Company has invested approximately $190 million and supported more than 50 proven methods, according to its website.[10] In addition, it provided three grants at $25,000 each in 2005 under a newly-focused program in this area. Colgate-Palmolive awarded grants in the range of $10,000-40,000 in 2005 and consistently funds post-doctoral fellowships at $33,500.[11] CTFA provided a three-year, $1 million per year grant to The Johns Hopkins School of Public

[1] Wayland, Susan H., Deputy Assistant Administrator, Environmental Protection Agency letter to participants in United States High Production Volume Challenge Program, October 14, 1999.
[2] Report Language in House VA, HUD and Independent Agencies Appropriations Bill for fiscal year 2003, H.R. 2620.
[3] Report Language in House VA, HUD and Independent Agencies Appropriations Bills for fiscal years 2004, 2005 and Report Language in House Interior Appropriations Bill for fiscal year 2006.
[4] http://thomas.loc.gov and http://niehs.nih.gov
[5] United States Public Law Number 106-545, http://thomas.loc.gov
[6] European Commission Directorate General Joint Research Centre Communication, March 25, 2003, Hartung, Thomas, Head of Unit, ECVAM correspondence on August 9, 2005.
[7] Hartung, Thomas, Head of Unit, ECVAM correspondence on August 9, 2005.
[8] www.bfr.bund.de/cd/1591
[9] www.nc3rs.org.uk/page.asp?id=30
[10] www.pg.com/science/ria_approach.jhtml
[11] http://altweb.jhsph.edu/databases/funding3.htm

Health Center for Alternatives to Animal Testing in 1981-1983. It also provided a grant to the same entity in 1984 at $700,000, with "continued funding in subsequent years."[12]

Private/public science Sector

Activity in this area in the science sector outside of the federal governments and industry has burgeoned worldwide in the past decades, including through the Institute for *InVitro* Sciences (IIVS), the Dr. Hadwen Trust for Humane Research, the Swedish Fund for Research Without Animal Experiments, the Fund for the Replacement of Animals in Medical Experiments (FRAME) and The Johns Hopkins School of Public Health Center for Alternatives to Animal Tests (CAAT). IIVS, which is based in the United States, has expended $10.3 million between 1997 and 2005.[13] This amount includes research and development and actually running methods to generate data to publish to boost the number of companies that use the methods. To date in 2005, the Dr. Hadwen Trust expended over 250,000 pounds.[14] FRAME's 2003-2004 annual report states that it spent 361,363 pounds on research. The Swedish Fund for Research Without Animal Experiments, a private fund also funded by public donations, awards grants each year totally approximately 1-1,5 million Euros.[15] Finally, in this brief analysis, the CAAT estimates it funded 300 grants between 1981-2005 for $5.5 million.[16]

Animal advocacy foundations or organisations

Non-profit organisations have played a significant role in funding non-animal and other alternatives development over the years. Animal organisations, largely funded by individual donors and subject to the scrutiny of non-profit reporting, are often more transparent than other funding sources.

People for the Ethical Treatment of Animals is funding numerous projects between 1998-2007 for approximately $590,000.[17] The American Fund for Alternatives to Animal Research, founded by Dr. Ethel Thurston who is a true pioneer in this area, has channeled $1 million into research and development since 1977.[18] The Humane Society of the United States, with more than 9 million supporters and constituents, provides a stipend of $5,000 every three years with its Russell and Burch Award.[19] The Alternatives Research Development Fund issues approximately $150,000 each year.[20] And the Doris Day Animal League provides up to $5,000 on an annual basis for special projects.

Discussion and conclusions

The lack of easily accessible and collated information from the various federal agencies or competent authorities hinders the overall ability to deliver a clear picture of the worldwide situation. In the United States alone, there are 15 federal regulatory and research agencies with a variety of individual programs that support both extramural and intramural research. While a few of the agencies have been influenced either internally or externally to set aside directed funding, it is imperative that an integrated, worldwide approach to prioritisation of endpoints, funding for research, development and validation of non-animal and other alternative test methods and a genuine push for their acceptance occur. It is vital to growing the science.

The private scientific sector is providing considerable funding for research and development of alternative methods, but due to their individual mandates and interests there is not a co-ordinated approach to focusing on specific endpoints or even areas of research. The same concern pertains to the animal protection organisations. While the organisations have a tendency for dialogue among themselves, often the projects to be funded are not the focus of discussion, rather the emerging science as a whole.

By the 6th World Congress on Alternatives and Animal Use in the Life Sciences, the various federal entities, industry, the private scientific sector and animal protection organisations should create a collective document clearly delineating the funding spent to research, develop, and validated non-animal and other alternative methods. There should be a direct attempt to track this funding from the point when a method demonstrates potential relevance to a specific endpoint, instead of lumping together all spending from very early stages of research.

Correspondence to
Sara Amundson
Deputy Director/Legislative Director
Doris Day Animal League
227 Massachusetts Ave., NE
Suite 100
Washington, DC 20002
USA
e-mail: info@ddal.org

[12] www.ctfa.org/Content/NavigationMenu/About_CTFA/History/History_4.htm
[13] Curren, Rodger, President of the Institute for InVitro Sciences correspondence on August 17, 2005.
[14] Langley, Gill, Scientific Advisor for Dr. Hadwen Trust for Humane Research on July 25, 2005 and www.drhadwentrust.org.uk
[15] www.stifud.se/print.php?filename=english/index.php
[16] http://altweb.jhsph.edu/databases/funding3.htm
[17] Seidle, Troy and Sweetland, Mary Beth, People for the Ethical Treatment of Animals correspondence on August 8, 2005.
[18] http://altweb.jhsph.edu/databases/funding3.htm
[19] www.hsus.org/ace/12630
[20] Leary, Sue, President of the Alternatives Research Development Fund in correspondence on August 8, 2005.

The OECD Health Effects Test Guidelines for REACH Need Updating

Robert D. Combes
FRAME, Russell & Burch House, Nottingham, UK

Summary
Each OECD Health Effects Test Guideline has been analysed for suitability to be used in the REACH system for chemicals testing. Some TGs are unnecessary while others urgently require updating to include modern techniques for improving the scientific and animal welfare aspects of the procedures involved. The reasons why the TGs have become so outdated are discussed and recommendations are made for improving this unsatisfactory situation. It is recommended that the OECD and its international experts focus more on revising outdated, and deleting obsolete, TGs.

Keywords: OECD test guidelines, REACH, revision, deletion, three Rs

Introduction

Background to REACH

The European Commission (EC) issued a White Paper in 2001 entitled *Strategy for a Future Chemicals Policy*, aimed at ensuring greater protection of human health and the environment (Anon., 2001; Combes, 2004a; Warbrick and Evans, 2004) and to harmonise the way in which risk assessments are conducted for new and existing chemicals. This policy requires risk assessments to be undertaken for existing and new chemicals, prior to their regulation according to a new registration and evaluation scheme, REACH (*R*egistration, *E*valutation and *A*uthorisation of *C*hemicals). The REACH legislation also shifts the burden of responsibility from regulatory bodies to industry to demonstrate the safety of chemicals. Industry will be required to submit detailed information on individual substances, including hazard property data, classification and labelling information, up to date safety sheets, and use and exposure information across the supply chain.

The European Commission issued its formal draft proposals for REACH which were adopted in October 2003, including the suggestion to create a new European Chemicals Agency to implement the REACH System in conjunction with the Competent Authorities (CAs) in Member States and the Commission Services (Dandrea and Combes, 2003). These proposals were issued in the form of an internet consultation and prompted many concerns on the part of industry and animal protection and welfare groups. They also contained a number of annexes, among which was one that included all the OECD Health Effects Test Guidelines, on the presumption that these would be required to be used for obtaining hazard information on chemicals under the REACH legislation.

The OECD Health Effects Test Guidelines

The OECD Test Guidelines (TGs), a collection of methods recommended to assess the hazards of chemicals, preparations and products, are intended to harmonise toxicity testing and to establish minimal standards. The OECD considers its TGs to be '*…the standard reference tool for chemical testing*' (preamble to TGs).

We have previously conducted a critical appraisal (Combes et al., 2004; Combes, 2004b) of the suitability of the TGs for REACH and found that several are in urgent need of revision with respect to techniques and replacement, reduction and refinement strategies (Balls et al., 1995). More information regarding these deficiencies and suggestions for improving the OECD TGs Programme are now presented.

Analysis of the OECD Health Effects TGs

Unnecessary TGs

Several TGs are not required for REACH (tab. 3 in Combes et al., 2004). Thus, TGs 401 and 406 (acute oral toxicity – LD_{50} method and skin sensitization, respectively) have been superseded by improved *in vivo* protocols. TG 401 has been withdrawn by the OECD and replaced by 420, 423 and 425E. TGs 404, 405, and 427 (the corrosion part of acute dermal irritation and corrosion, acute dermal photoirritation, and dermal absorption, respectively) have been replaced by validated and endorsed *in vitro* methods. 415 and 476-485 (one generation reproduction study and various genotoxicity assays, respectively) are not required scientifically. Use of 424 (neurotoxicity in rodents) would be expected to cause considerable stress to animals, and also the end point is included in other TGs.

The OECD regards 428 (skin absorption: *in vitro* method) as being a full replacement for 427 (the *in vivo* method for skin absorption) (tab. 1), thus a TG for this endpoint *in vivo* is not needed. As the OECD views 432 as a complete replacement method, an *in vivo* phototoxicity protocol is also unnecessary.

TGs needing updating

Five TGs are clearly in urgent need of revision, four of which were originally adopted in 1981 and the remaining one in 1984, with none having been updated in the intervening period (tab. 2).

Apart from basic animal welfare issues, such as group housing and environmental enrichment, more worrying is the fact that these TGs ignore substantial advances made to the collection and analysis of information, and in the use of humane endpoints. This is particularly the case for 417 and 451 (toxicokinetics and carcinogenicity, respectively). Such techniques include non-invasive imaging (Hudson, 2005), better histopathology (Chejfee, 1999), humane endpoints (Combes et al., 2002) and biomarkers (Bottrill, 1998) (tab. 3).

The OECD lists 15 updated TGs (Anon., 2005a). However, the number of updated TGs is in reality less than the total listed (tab. 4). Thus, for 408, 409, 416, 418 and 419 no details of revisions are given, except for a year. Moreover, 420, 423 and 425 similarly have years for revision, but all supersede 401 (Botham, 2002), and not previous respective guidelines. Revisions of seven TGs do involve increased implementation of reduction and refinement (tab. 4), however.

Tab. 1: An analysis of the OECD *in vitro* test guidelines relevant to REACH

TG	Title	OECD notes	Comments
428	Skin absorption: *in vitro* method	Full replacement for TG 427 (*in vivo* method)	Therefore, TG 427 should be deleted
430	*In vitro* skin corrosion: TER	Replacement for corrosion part of TG 404	Therefore, TG 404 needs updating
431	*In vitro* skin corrosion: human skin model test	Ditto	Ditto
432	*In vitro* 3T3 NRU phototoxicity test	Full replacement (no existing TG)	Therefore, no need to develop an *in vivo* TG
435	*In vitro* skin corrosivity	Replacement for corrosion part of TG 404 for acids and bases only	Therefore, TG 404 needs updating; is there a need for both TGs 430 and 435?

Tab. 2: Some OECD Health Effects Test Guidelines that need updating

TG number (date)	Title	Comments
412 (1981)	Repeated dose inhalation toxicity: 28d or 14d study	Stressful procedures could be minimised by preliminary dose range finding study; blood collection before dosing should be avoided; dosing regimen, tissue retention and analysis could all be improved, effects of head only exposures on eyes needed (ophthalmoscopy)
417 (1984)	Toxicokinetics	Methods should involve: modern technology; guidance on housing and re-use of animals; prior hepatocyte screening for species selection; guidance on the need for PBPK modelling and full ADME information:
451 (1981)	Carcinogenicity studies	Revision to take account of new technical developments for tumour identification and characterisation; humane endpoints (including non-invasive imaging), timing of studies; observation of animals
452 (1981)	Chronic toxicity	Revision to clarify: group sizes; need for test when 90d study done; dosing regimens and routes; specific justification for second species
453 (1981)	Combined chronic toxicity/ carcinogenicity studies	Revision required to address: why chronic element only applies to high dose and control groups; mode of administration and dosing regimens; use of dietary restriction; number of species

Tab. 3: Some modern techniques that could potentially improve the OECD Health Effects Test Guidelines both scientifically and with respect to animal welfare

Type of technique	Methods available
Molecular biomarkers	Toxicogenomics; proteomics; metabonomics
Non-invasive imaging/labelling	Bioluminescence; fluorescent markers; quantum dots; PET imaging; MRI scanning; ECG ventral plane videography (e.g. for early i.d. of tumours)
Body fluid analysis	AMS (accelerated mass spectrometry); NMR spectroscopy; PET imaging; identification of macromolecular adducts
Refined procedures (to facilitate compound administration; blood collection; measurement of blood pressure, heart rate)	Indwelling intravenous catheters; implanted vascular access ports; implanted telemetry devices
Refined procedures relating to pathological examination	Non-terminal and minimally invasive methods for obtaining body fluid and tissue samples – e.g. metabolic cages where excretia can be collected for analysis; immuno-histochemical analysis; confocal microscopy; fluorescence in situ hybridisation (FISH); PCR; PAGE; detection of apoptotic cells; differential screening of phage displayed libraries to identify cell surface markers

Integrating the use of the OECD TGs with non-animal methods

Some of the above improvements (tab. 4) have been achieved by using the TGs in a tiered testing strategy with *in vitro* assays. Combes et al. (2004) proposed the use of 11 non-animal methods with the OECD TGs in an integrated testing approach for REACH, such that the more animal-intensive TGs would not have to be used in many cases where positive data were obtained in earlier tests (tab. 6 in Combes et al., 2004; fig. 1; Combes et al., 2003; Combes et al., 2004; Grindon et al., 2005).

The above 11 methods are either awaiting formal validation (e.g. (Q)SAR, expert system and PBPK modelling and the SHE cell transformation assay), or are being validated (e.g. the use of basal cell cytotoxicity to predict acute toxicity), or have been validated (e.g. screens for eye irritation and embryotoxicity) (Spielmann and Liebsch, 2002).

The need for the OECD to focus more on updating TGs

The current levels of revision

The OECD has revised several TGs to reduce group sizes, change dosing regimes, and obtain more information from each animal (tab. 2 and 4). However, often technical changes to endpoint scoring have lagged behind new technologies.

Since 1981, 41 new guidelines have been adopted, one of which has been deleted, and 16 have allegedly been updated, although 8 of these were not true updates (tab. 5). This implies that the OECD places more emphasis on producing new TGs than on revising or deleting TGs.

The OECD system for producing new TGs

The OECD has schemes for: a) defining the need for a TG; and b) producing new TGs (Koeter and Visser, 2000). Neither of

Tab. 4: OECD analysis of Health Effects Test Guidelines (animal tests) relevant to REACH that have been updated/deleted

TG number	Title	OECD notes on date & details of updating	Comments
401	Acute oral toxicity	Date of deletion: 20 Dec. 2002	Other TGs rely on death as an endpoint
402	Acute dermal toxicity	24 Feb. 1987; reduction compared to original TG, lowering of dose level	
404	Acute dermal irritation/ corrosion	24 April 2002; refinement/reduction – part of tiered testing strategy with *in vitro* screens	Corrosivity part should not be required as replacement test available (TG 430)
405	Acute eye irritation/ corrosion	Ditto	Ditto
406	Skin sensitisation	17 July 1992; reduction method – by 50% compared to original TG	Should be deleted or guidance produced for its use only when LLNA unsuitabl
407	Repeated dose 28d oral toxicity in rodents	27 July 1995; refined guideline with more information on: best dosing practice and from each animal	
408	Repeated dose 90d oral toxicity in rodents	21 Sept. 1998: animal test	Details of the updating not presented
409	Repeated dose 28d dermal toxicity in rodents	Ditto	Ditto
414	Prenatal developmental toxicity	22 Jan. 2001; reduction of 20% in number of animals used and more information from same animal	
416	Two generation reproductive toxicity	22 Jan. 2001; animal test	Details of the updating not presented
418	Delayed neurotoxicity of organophosphorus substances following acute exposure	27 July 1995; animal test	Ditto
419	Delayed neurotoxicity of organophosphorus substances; 28d repeated dose	Ditto	Ditto
420	Acute oral toxicity Fixed Dose Procedure	17 Dec. 2001; reduction and refinement of TG 401 (less suffering; smaller number of animals)	Not an update of the TG 420, but of 401
423	Acute oral toxicity – Acute Toxic Class method	Ditto; much smaller number of animals (10% of TG 401)	Ditto; still relies on death as endpoint
425	Acute oral toxicity – Up and down method	Ditto; provides a closer estimate of LD_{50} than TGs 420 and 423	Ditto; still relies on death as endpoint

these schemes includes any overt mention of specific mechanisms for updating or deleting TGs, although this is probably done via the existing system. A formal mechanism would, however, encourage the OECD to pay more attention to the need for these processes.

TGs for animal testing methods have traditionally become established and 'validated' following many years of use. Then, in the 1980s, several genotoxicity assays were validated by different interlaboratory validation studies before the protocols became OECD TGs. However, more recently, several refined *in vivo* and *in vitro* TGs have been produced following more stringent validation studies and peer reviews, according to internationally-agreed criteria, involving also the OECD (OECD, 2003; NIH, 1997; Combes, 2003).

The current focus of the OECD on producing new TGs

There are inconsistencies in the approach that the OECD has for developing TGs. Thus, it plans to introduce eight new TGs (tab. 6) under different circumstances. TG numbers have been assigned to, and drafts written for, TGs for inhalation and dermal exposure by using the Fixed Dose Procedure (FDP) and one for *in vivo* skin absorption test); all when there have apparently been no formal independent and totally transparent validation studies and peer reviews. A validation study of the FDP tests could probably be achieved by a retrospective weight of evidence review (Balls and Combes, 2005a). This is because the FDP for acute oral toxicity has been validated, although it took many years for a TG to be written (Combes et al., 2002). The *in vitro* micronucleus assay has been assigned a TG number and a draft guideline has been produced (OECD, 2004), following interlaboratory studies (Kirsch-Volders, 1997). It is being validated retrospectively by ECVAM, then peer reviewed by the ESAC.

Some other test methods, either not validated or lacking a peer review report, are also being considered for TG development, e.g. the uterotrophic assay, the validation study of which was flawed (Combes, 2003), and the peer-review of which was not independent (Combes, 2004c), the OECD is even proceeding to draft a TG without a final published peer-review. Also, detailed review papers (DRPs) for the use of transgenic mice in carcinogenicity testing and the SHE cell transformation assay are being written, as a prelude to TGs, the former despite the failure of an interlaboratory study (van Zeller and Combes, 1999; Committee on Carcinogenicity, 2002).

Clearly, several of these tests are unnecessary or unready for conversion into TGs, and reflect the OECD's over-emphasis on guideline production. Several regulatory authorities are also prematurely accepting these methods (tab. 6), which is scientifically unjustified and could result in unreliable hazard data being produced.

Discussion

Strategies for test development, validation, peer review, TG production and adoption and regulatory acceptance are clearly inconsistent and require overhaul (as shown in fig. 2). The original approach (scheme I; fig. 2) was developed when *in vivo*

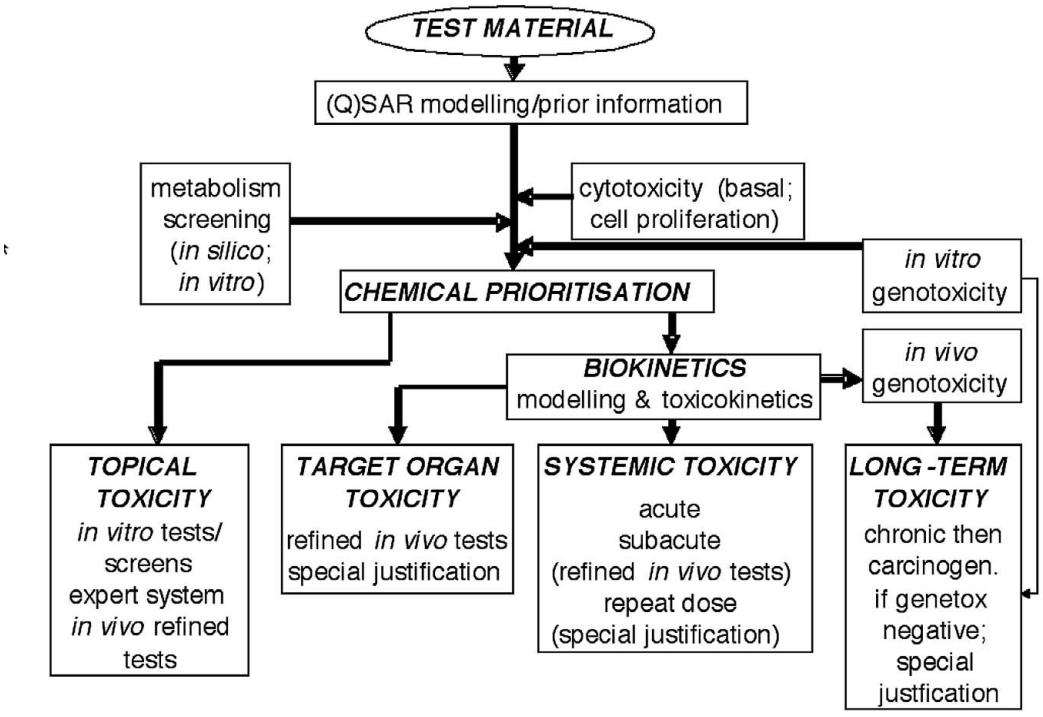


Fig. 1: Scheme for using non-animal and animal methods, including OECD Test Guidelines, in an integrated testing scheme for REACH (modified from Combes et al., 2004).

Tab. 5: OECD analysis of Health Effects Test Guidelines (animal tests) relevant to REACH

Date originally adopted/updated/deleted	Test Guidelines	Total number
1981	401, 402, 403, 404, 405, 406, 407, 408, 409, 410, 411, 412, 413, 414, 451, 452, 453	17
1983	415, 416	2
1984	417, 418, 419, 478	4
1986	484, 485	2
1987	401, *402, 406*	1 *(2)*
1992	*404*, 420	1 *(1)*
1995	*407, 418, 419*, 421, Acute dermal photo-irritation screening (draft), 422	3 *(3)*
1996	*403*, 423	1 *(1)*
1997	424, 474, 475, 483, 486	5
1998	*408, 409*, 425	1 *(2)*
2000	426 (draft new)	1
2001	*414, 416, 402E, 423E, 425E*, 427 (draft)	1 *(5)*
2002	**401**, *404E*, *405E*, 429E	1 *(2)* **(1)**
2004	427	1
Totals		41 *(16)* **(1)**

*updated tests**, **deleted tests**

**8 of these are unspecified or not true updates of the respective TGs*

Tab. 6: Some test methods recently considered by the OECD for test guideline development.

Test	Status*					
	DRP	Validation study completed	Peer review endorsement	Draft TG	Regulatory approval	Notes
TG 433 FDP inhalation	Y	N	N	Y	N	No formal validation study, but method assessed by experts
TG 434 FDP dermal	Y?	N	N	Y	N	Ditto
TG 435 *in vitro* corrosivity	Y?	Y	Y	Y	Y	Specific for acids and bases
TG 427 *in vitro* skin absorption	?	N	N	Y	N	Limited scientific review; no scientific justification for TG
In vitro micronucleus	?	(Y)	N	Y TG487?	(Y)	Several small validation studies, results being subjected to a weight of evidence validation by ECVAM; ESAC to peer-review validation; OECD issued draft TG on 14.06.04; UK HSE accepts data
SHE cell transformation assay	Y	N	N	N	(Y)	ECVAM planning validation study; FDA accepts data to clarify other information
Uterotrophic assay	(Y)	Y	(Y)	(Y)	(N)	Published paper serves as DRP; peer review organised by OECD, but final report not been published; OECD proceeding with draft TG; Japanese authorities already accepting data from test
Transgenic mouse carcinogenicity assays	(Y)	(Y)	N	N	(Y)	DRP in preparation, based on unsuccessful ILSI validation; FDA encourages/accepts data from assays.

* parentheses indicate that answer is qualified

methods were never formally validated. Later, some tests, primarily those intended as non-animal replacements, or refinements of tests, underwent intensive validations and peer reviews before TGs production (scheme II; fig. 2). In contrast, other tests have not been properly validated and peer reviewed, yet have, or will have, TGs (scheme III; fig. 2). Also, premature regulatory acceptance has occurred (scheme IV; fig. 2)! A combination of all schemes (fig. 2) is in use in what seems to be an arbitrary way, probably as some consider *in vivo* tests to be inherently the most reliable, being based on intact animals, a scientifically absurd generalisation (Balls, 2004). Adoption of scheme II (Fig 2) would avoid premature TG development and regulatory acceptance.

The OECD is evaluating its Test Guidelines programme, and released an issues paper on 16 March, 2005 (Wagner, 2005; "Refocus of the Test Guidelines Programme") for discussion at the NCs meeting the following month. The paper, referring to Combes et al. (2004), stated: "*Updating these test Guidelines could be approached in different ways. Refinements and improvements could be made 'within' the test, to ensure that the test parameters and protocols reflected current science and regulatory requirements ….or there could be improvement and enhancement 'across' tests, where test methods are modified and improved on the understanding of the role that the method plays in testing.*" However, unfortunately at the meeting (Anon., 2005b) updating of the TGs programme as a whole was apparently not discussed.

The need to validate revised TGs before their adoption should be considered case by case, and should utilise weight of evidence review as far as possible. A requirement for practical validation should not, however, detract from having the most modern TGs available.

Conclusions and recommendations

The OECD TGs programme is unsatisfactory; lacking a formal system for revising and deleting TGs, and needs overhaul. The OECD should just produce, update and delete redundant and obsolete TGs, instead of validating tests and producing guidance on validation (OECD, 2003) that it does not follow itself. The TGs might then become state of the art methods and more suitable for REACH.

Other bodies, like ECVAM, ICCVAM and JECVAM, should manage validation studies. No TG should be prepared, and no test accepted by a regulatory body, before satisfactory validation and peer review. If these processes are unsatisfactory the test should be considered for invalidation (Balls and Combes, 2005b).

Hopefully, the OECD will have the foresight to grasp this opportunity to revise the Test Guidelines Programme in a way that brings about these necessary changes.

References

Anon. (2001). *White paper on a Strategy for a Future Chemicals Policy*, COM(2001) 88 final, 32pp. Brussels, Belgium: Commission of the European Communities. Website http://europa.eu.int/eur-lex/en/com/wpr/2001/ com2001_0088en01.pdf (Accessed 6.10.04).

Anon. (2005a). *OECD Guidelines for testing of chemicals, Appendix 4*. http://pharmacos.eudra.org/F3/cosmetic/doc/AnTest/(6)%20Appendixes/Appendix%20IV.pdf

Anon. (2005b). Test Guidelines Programme – Draft Summary Record of the 17th Meeting of the Working Group of National

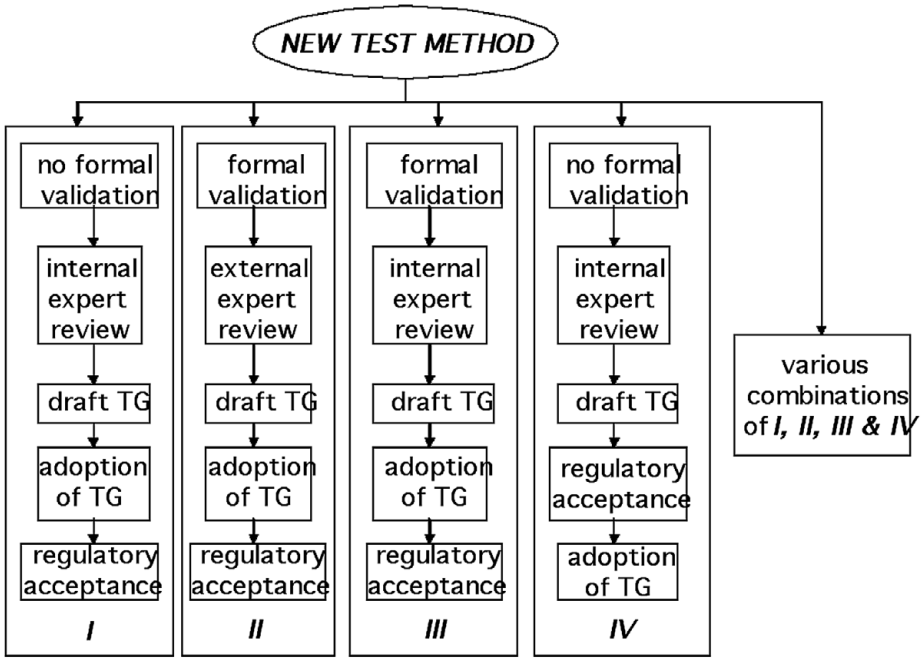


Fig. 2: A summary of the various schemes for the development and adoption of OECD Test Guidelines and regulatory acceptance.

Co-ordinators of the Test Guidelines Programme (WNT), 12-14 April, 2005, *OECD Paris, ENV/JM/TG(2005)1*, pp. 31.

Balls, M. (2004). Are animal tests inherently valid? *ATLA 32 (Suppl.1)*, 755-758.

Balls, M. and Combes, R. (2005a). Validation via weight of evidence approaches. *ALTEX* (this volume).

Balls, M. and Combes, R. (2005b). The need for a formal invalidation process for animal and non-animal tests. *ATLA 33*, 299-308.

Balls, M., Goldberg, A. M., Fentem, J. H. et al. (1995). The Three Rs: The way forward. The report and recommendations of ECVAM workshop 11. *ATLA 23*, 838-866.

Botham, P. A. (2002). Acute systemic toxicity. *ILAR Journal 43*, S27-S30.

Bottrill, K. (1998). The use of biomarkers as alternatives to current animal tests on food chemicals. *ATLA 26*, 421-480.

Chejfee, G. (1999). *Recent Advances in Histopathology*. New York: Churchill-Livingstone Inc.

Combes, R. D. (2004a). The current OECD Health Effects test Guidelines are in urgent need of revision. *ATLA 32*, 463-464.

Combes, R. D. (2004b). The EU REACH scheme for chemicals testing – a challenge and an opportunity for alternatives to laboratory animals. *Science in Parliament, Summer 2004*, 8-9.

Combes, R. D. (2004c). Peer review of validation studies: an assessment of the role of the OECD by reference to the validation of the uterotrophic assay for endocrine disruptors. *ATLA 32*, 111-117.

Combes, R. D. (2003). A critical assessment of the OECD collaborative study to validate the uterotrophic assay for the detection of oestrogenic and anti-oestrogenic chemicals. *ATLA 31*, 489-499.

Combes, R., Schechtman, L., Stokes, W. S. and Blakey, D. (2002). The international symposium on regulatory testing and animal welfare: recommendations on best scientific practices for subchronic/chronic toxicity and carcinogenicity testing. *ILAR Journal 43 (Suppl.)*, 112-117.

Combes, R. D., Gaunt, I. and Balls, M. (2004). A scientific and animal welfare assessment of the OECD health effects test guidelines for the safety testing of chemicals under the European Union REACH system. *ATLA 32*, 163-208.

Combes, R., Barratt, M. and Balls, M. (2003). An overall strategy for the testing of chemicals for human hazard and risk assessment under the EU REACH system. *ATLA 31,* 7-19.

Combes, R., Balls, M., Bansil, L. et al. (2002). An assessment of progress in the use of alternatives in toxicity testing since the publication of the report of the Second FRAME Toxicity Committee (1991). *ATLA 30*, 365-406.

Committee on Carcinogenicity (2002). *COC statement on ILSI/HESI research programme on alternative cancer models COC/02/S3*. http://www.advisorybodies.doh.gov.uk/ coc/ ilsihesiact.htm.

Dandrea, J. and Combes, R. D. (2003). A survey of stakeholder organizations on the proposed new European chemicals policy. *ATLA 31*, 501-528.

Grindon, C., Bhogal, N., Combes, R. and Balls, M. (2005). The scientific, animal welfare and regulatory implications of the EU REACH system for chemical safety. *TOPRA Regulatory Rapporteur 2, No. 4*, 16-22.

Hudson, M. (2005). The welfare and scientific advantages of non-invasive imaging of animals used in biomedical research. *Animal Welfare 14,* 303-317.

Kirsch-Volders, M. (1997). Towards a validation of the micronucleus test. *Mutation Res. 392*, 1-4.

Koeter, H. B. W. M. and Visser, R. (2000). Work in OECD on chemical safety: approaches for human risk assessment. *Industrial Health 38*, 109-119.

NIH (1997). *Validation and Regulatory Acceptance of Toxicological Test Methods. A Report of the ad hoc Interagency Coordinating Committee on the Validation of Alternative Methods*. Research Triangle Park, NC, USA: NIEHS.

OECD (2003). *Draft Guidance Document No 34 on the Development, Validation and Regulatory Acceptance of New and Updated Internationally Acceptable Test Methods in Hazard Assessment, Environment Directorate*. Paris, France: OECD.

OECD (2004). *OECD guideline for the testing of chemicals: draft proposal for a new guideline 487: In vitro micronucleus test. Draft guideline, June 14, 2004* (1st version). Paris, France: OECD.

Spielmann, H. and Liebsch, M. (2002). Validation successes: chemicals. *ATLA 30 (Suppl. 2)*, 33-40.

van Zeller, A-M. and Combes, R. D. (1999). Transgenic mouse bioassays for carcinogenicity testing: a step in the right direction? *ATLA 27 (Suppl. 1)*, 839-846.

Wagner, D. (2005). Test Guidelines Programme – Issues Paper: Refocus of the Test Guidelines Programme, 17th Meeting of the Working Group of National Co-ordinators of the Test Guidelines Programme (WNT), 16 March, 2005, OECD Paris, *ENV/JM/TG(2005)*13 pp. 23.

Warbrick, E. V. and Evans, P. F. (2004). The regulatory acceptance of alternatives in the European Union. *ATLA 32, (Suppl. 1)*, 657-661.

Acknowledgement

The author thanks Michael Balls and Ian Gaunt for useful discussions.

Correspondence to

Prof. Robert C. Combes
FRAME
Russell & Burch House
96-98 North Sherwood Street
Nottingham NG1 4EE, UK
e-mail: Bob@frame.org.uk

Challenges and Opportunities of Animal Welfare Organisations in Influencing and Making Public Policy

Ursula G. Sauer
Akademie für Tierschutz, Neubiberg, Germany

Summary
For ethical and scientific reasons, the German Animal Welfare Federation strives for an end to all animal experiments. Amongst other issues, its political activities lately have been dedicated campaigns to end the use of non-human primates in research and on the new EU Chemicals Policy, REACH. Animal welfare requests are backed up by strong societal support. In influencing public policy animal welfare issues have to be weighed against other societal concerns, such as human health and environmental protection, and personal or economic interests. From the point of view of animal welfare, animal experimentation can be stopped without impeding human health or environmental protection.

Keywords: animal welfare, political activities, social values, non-human primate research, ethical evaluation, EU chemicals policy, REACH, non-animal testing strategies, animal testing ban

Introduction

According to Thomas Dye (2001), public policy is "*whatever governments choose to do or not do*." As Schiffelers et al. (2005) note, a large part of the political activities of governments is aimed at preparing and implementing legislation: "*The development of legislation in the EU is a complex process in which many institutions, actors and (formal and informal) rules play a role. By far the largest part of EU policy – and this applies to animal testing as well – eventually manifests itself in European law (regulations and directives). Within this formal, institutional framework, various actors use all kinds of influence and pressure with the aim of obtaining the desired legislative results. At the European level, lobbying is part and parcel of the legislative process. In many cases, consultation with special interest groups is a fixture in the formulation of policy.*"

In the process of influencing and making public policy, animal welfare organisations see it as their role to speak up for the welfare of the individual animal. In accordance with Council Directive 86/609/EEC on the protection of laboratory animals (Commission of the European Communities, 1986), an "*experimental animal*" is being used for a scientific procedure, "*which may cause it pain, suffering, distress or lasting harm*". Therefore the final goal in striving to protect these animals can only be the total abolition or replacement of such experimentation.

At the same time, animal welfare organisations acknowledge the contribution that reduction and refinement methods, the other two Rs of the so-called 3Rs principle (Russell and Burch, 1959), make in improving the welfare of laboratory animals. Reduction implies that the new methodology uses fewer laboratory animals to pursue the same goal as before, whereas refinement means that the distress of the animals in a given procedure is reduced by changing the methodology. From the point of view of animal welfare such improvements are to be welcomed, however they can only be considered interim achievements.

Animal welfare organisations are backed up by strong societal support. The EU citizens' concern about the welfare of animals is also reflected in the fact that animal welfare has been included in the Constitutional Treaty (Anon, 2004). Article III-121 reads: "*In formulating and implementing the Union's … policies, the Union and the Member States shall pay full regard to the welfare requirements of animals, as sentient beings…*". A European survey from June 2005 on "Social values, Science and Technology"[1] came to the conclusion: "*82% of EU citizens uphold our duty to protect the rights of animals whatever the cost… in twenty-seven of the surveyed countries at least three in four share this point-of-view.* "

While animal welfare organisations engage in political activities to improve the welfare of animals, other stakeholders representing for example the scientific community, industry or environmental and consumer protection organisations, also put forward their requests. Thus, when advocating in favour of animal welfare issues, the following considerations have to be taken into account: Do animal welfare issues compete with the issues of human health and environmental protection? Do animal welfare issues compete with other issues, such as economic or personal interests?

These questions can only be answered on a case by case basis. In order to determine the significance and the true implications of any argument, either put forward against or in favour of animal testing, the underlying intention should be revealed and the scientific evidence brought forward to back it up should be evaluated. Again and again it can be shown that not only ethical, but

[1] Special Eurobarometer 225: Social values, Science and Technology, June 2005; http://europa.eu.int/comm/public_opinion/archives/ebs/ebs_225_report_en.pdf

also scientific reasons speak against animal experiments and that turning to non-animal test methods does not impede human health or environmental protection.

In the following, two examples are presented to show which challenges and opportunities animal welfare organisation face when striving for an end to animal experimentation: The campaign to ban research with non-human primates and the campaign to prevent animal testing under the new EU-Chemicals Policy.

Campaign to end research with non-human primates

While animal welfare organisations strive for an end to all experimentation with sentient animals, research with non-human primates always has been in the focus of their attention. Since these animal species are man's next of kin in the animal kingdom, using them in procedures that would be considered unacceptable in humans inevitably leads to an ethical conflict that cannot be disregarded. Every year, approximately 10,000 non-human primates are being used for experimental purposes in the European Union, and a reduction of this number is not discernible. Non-human primates are being used in fundamental research, for example in neurological or immunological studies, in an attempt to use an animal species that is as closely related to human beings as possible. For this same reason they are also being used in applied research, such as in toxicokinetic or pharmacodynamic studies. In a scientific and ethical evaluation of experiments with non-human primates, Ruhdel and Sauer (1998) revealed that the animals not only suffer from the procedures, but also due to acquisition and housing, whereas outstanding medical benefits that would justify the use of these animals could not be identified. From the point of view of animal welfare, there are compelling ethical reasons to discontinue experiments with non-human primates, while a multitude of adequate non-animal test methods are available to pursue the scientific questions currently addressed in experiments with these animals (Ruhdel and Sauer, 1998).

The European Commission has acknowledged the special concern related to the use of non-human primates in research. In the Preamble to Council Decision 1999/575 (Commission of the European Communities, 1999), in which the European Convention ETS 123 for the protection of laboratory animals (Council of Europe, 1986) was approved, it is stated: "*Whereas the use of primates for experimental and other scientific purposes carries the risk of suffering for those animals and therefore has to be reduced; whereas the use of primates for experimental and other scientific purposes has led to the catching of primates in the wild, and whereas, this should be avoided whenever possible in view of the suffering and losses which can arise during catching and transport.*"

The European Commission also specifically mentions the special concern for non-human primates and their inadequate covering by the current Directive to justify the need for the up-coming revision of Council Directive 86/609 on the protection of laboratory animals: "*Nor is the use of animals with a higher degree of neurophysiological sensitivity specifically regulated, such as in the case of non-human primates.*"[2]

On the level of the Council of Europe, legal activities relating to the scientific use of non human primates recently took place in the context of the revision of Appendix A "*Guidelines for accommodation and care of animals*" of European Convention ETS 123 on the protection of laboratory animals (Council of Europe, 1986). After the Parties to the Convention launched this revision at the 3rd Multilateral Consultation in May 1997, species specific expert groups were formed to design species specific guidelines for the housing of research animals. These experts were nominated from the non-governmental organisations, which attend the meetings of the Parties as observers.[3]

A scientist from the German Animal Welfare Federation joined the Primate Expert Group as representative of the European umbrella organisation Eurogroup for Animal Welfare. Notwithstanding the request for a ban on all experiments with non-human primates, this engagement was considered relevant because of the great suffering these animals endure under standard housing conditions. The current minimum standards, which continue to be the rule for the majority of non-human primates kept in European laboratories, are totally inadequate to enable these animals to fulfill even the most basic of their ethological needs, such to live in stable harmonious groups, to engage in meaningful activities and to perform a normal repertoire of species-specific locomotor activity. It is still legal to keep non-human primates singly in small barren cages, in which they cannot even move a few steps in one direction, let alone sit on a perch with their head upright and the tail hanging freely.

Nevertheless the negotiations surrounding the revision of guidelines for the accommodation of non-human primates turned out to be difficult. There was strong opposition against changing the still valid minimum standards at all. From the point of view of animal welfare there is no justification if animals that are being used and sacrificed for the alleged benefit of humans are not even housed in a way that they can fulfill their very basic species-specific needs. So-called "scientific" arguments also cannot be put forward to reject this: Housing non-human primates in enriched group enclosures does not stand in the way to regularly handling these animals, if trained adequately (Reinhardt, 2002). Clearly the reluctance to change cage dimensions is neither driven by the wish to improve human health or environmental protection, but by personal or financial interests.

It is to be welcomed that the Parties to the Convention supported the requests of animal welfare. The final draft guidance document approved by the Parties and published in the internet[4] (Council of Europe, 2004) still is a compromise between economic and animal welfare issues, but it can make a considerable contribution to reducing the suffering of non-human primates due to inadequate housing conditions.

On the level of the European Union, animal welfare organisations will request a total ban on the use of non-human primates

[2] http://www.europa.eu.int/comm/environment/chemicals/lab_animals/
[3] http://www.coe.int/T/E/Legal_affairs/Legal_co-operation/
[4] For technical reasons, the official vote on the revised Appendix A is still outstanding. It currently is scheduled for the beginning of 2006.

for scientific purposes during the revision of Council Directive 86/609/EEC. In any case bans on the use of wild-caught non-human primates, on the genetic manipulation of non-human primates, and on procedures with medium or severe distress should be implemented as well as a ban on single housing of non-human primates except for limited duration on veterinary grounds.

EU chemicals policy campaign

In October 2003, the European Commission put forward the Draft REACH[5] Regulation (Commission of the European Communities, 2003). The aim of this new EU Chemicals Policy system is better to protect humans and the environment from unwanted effects of chemicals, while at the same time taking into account economic issues.

Animal welfare organisations have given the new EU Chemicals Policy high priority for campaigning, because of its risk of leading to a huge increase in animal testing – up to 45 million additional animal tests according to official predictions (Höfer et al., 2004). This should be prevented, also in the sake of human health and environmental protection. The redesign of the entire area of chemicals legislation should be used as a chance for a paradigm change in favour of non-animal safety testing strategies.

In order to ensure that REACH will truly serve to improve safety protection, flexible testing strategies should be developed and implemented for all relevant endpoints that make use of significant, scientifically validated non-animal test methods instead of unreliable animal tests. Nevertheless not even all currently available non-animal test methods have been included in the Draft Regulation. Schiffelers et al. (2005) consider resistance from authorities to be one of the main reasons for such problems: "*…any cost/benefit assessment usually favours an existing policy rather than a new one. New policy is often seen as a potential liability. This point is illustrated by a civil servant's comment on policy changes: 'It's better not to change ten times, than to make nine changes for the better and one for the worse'…Still, the reluctance of policy-makers and decision-makers to incorporate alternatives seems to result from a combination of factors such as attitude, knowledge and risk acceptance*." From the point of view of animal welfare, it is totally unacceptable if non-animal test methods that have shown their scientific validity according to internationally agreed criteria – a procedure that the vast majority of animal tests never were submitted to – are not included in regulatory testing strategies.

Animal welfare organisations also request data sharing to become mandatory in the REACH Regulation. The majority of the chemicals that will be covered under REACH are existing substances that have been in use for over 25 years. It is entirely impossible that there should be no data on their effects. They only never were reported to the authorities. Nevertheless, again and again representatives from industry speak up that they would prefer to pay a fee in order to be able to keep data to themselves. From the point of view of animal welfare such requests are unacceptable. They do not serve human health or environmental protection and clearly stand in the way to preventing animal testing.

All in all, in the future EU Chemicals Policy, full consideration of animal welfare issues is the key to ensure that the goals to improve human health and environmental protection can be met and they do take into account economic issues. However, to reach these goals, substantial amendments to the Draft REACH Regulation remain indispensable.

The EU legislative procedure is complex with a multitude of key players. The European Parliament and the Council are main key players since they have co-decision power on finalising the REACH Regulation. In the EU Parliament, legislative work is prepared in the Committees. Concerning REACH, the Environmental Committee has taken the lead, with Industry and Internal Market Committees also putting forward proposals for amendment and the Committees for Legal Affairs, Employment, Women's Rights, Economy, Budget and International Trade giving their opinions. The final proposals for amendment are adopted in the plenary of the Parliament.

In parallel the Draft REACH regulation is dealt with in the Environment and Competitiveness Councils. Preparatory work for the Council meetings is accomplished on the level of the Committees of Permanent Representatives, who formed Ad hoc Working Groups for this task.

The Council being formed by the Ministers of the Member States governments, its position is influenced by opinions formed on the national levels. The German Government, for instance, must take into account comments from the Upper and the Lower Chambers[6] and receives advice from three different authorities, the German Federal Institute for Risk Assessment (BfR[7]), the German Environmental Protection Agency (UBA[8]) and the Federal Institute for Occupational Safety and Health (BAuA[9]).

In influencing the development of the new Chemicals Policy, the work load of addressing all these key players is distributed between national animal welfare organisations and their European umbrella organisations, Eurogroup for Animal Welfare[10] and the European Coalition to End Animal Experiments.[11]

Conclusions

What would be the consequences of stopping animal testing as requested by animal welfare organisations in their engagements to influence public policy? Animal experiments are merely one type of test method amongst a multitude of others, and oftentimes scientific arguments can be put forward to show that they are unreliable or misleading. Research without animals seems possible, provided that scientists and politicians are willing to

[5] REACH – Registration, Evaluation, Authorisation of Chemicals
[6] Bundesrat, Bundestag
[7] BfR – Bundesinstitut für Risikobewertung
[8] UBA - Umweltbundesamt
[9] BAuA – Bundesanstalt für Arbeitsschutz und Arbeitsmedizin
[10] www.eurogroupforanimalwelfare.org
[11] www.eceae.org

change their attitudes, to develop new ways of thinking and to pursue new research strategies, by asking new scientific questions. As long as animal experiments are being performed, it is inevitable to address the question whether everything should be considered acceptable that is feasible. In pursuing the goal to prevent animal testing, solutions can be found that are compatible with other societal interests, such as improving human health and environmental protection. However, from the point of view of animal welfare, economic issues and personal interests should not be allowed to override animal welfare issues.

References

Note: all websites were accessed in October 2005.

Anon. (2004). Treaty establishing a Constitution for Europe. OJ C310, Volume 47. 16 December 2004. http://europa.eu.int/eur-lex/lex/JOHtml.do?uri=OJ:C:2004:310:SOM:EN:HTML.

Commission of the European Communities (1986): Council Directive 86/609/EEC of 24 November 1986 on the approximation of laws, regulations and administrative provisions of the Member States regarding the protection of animals used for experimental and other scientific purposes. OJ L 358, 18.12.1986, p.1-28 (http://europa.eu.int/smartapi/cgi/sga_doc?smartapi!celexapi!prod!CELEXnumdoc&lg=EN&numdoc=31986L0609&model=guichett).

Commission of the European Communities (1999). Council Decision 1999/575/EC of 23 March 1998 concerning the conclusion by the Community of the European Convention for the protection of vertebrate animals used for experimental and other scientific purposes. OJ L 222, p. 29-30 (http://europa.eu.int/eur-lex/pri/en/oj/dat/1999/l_222/l_22219990824en00290037.pdf).

Commission of the European Communities (2003). Proposal for a Regulation of the European Parliament and of the Council concerning the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH), establishing a European Chemicals Agency and amending Directive 1999/45/EC and Regulation (EC) {on Persistent Organic Pollutants} Proposal for a Directive of the European Parliament and of the Council amending Council Directive 67/548/EEC in order to adapt it to Regulation (EC) of the European Parliament and of the Council concerning the registration, evaluation, authorisation and restriction of chemicals. COM (003) 44 (03) (http://europa.eu.int/eur-lex/en/com/pdf/2003/com2003_0644en.html).

Council of Europe (1986). European Convention ETS 123 on the Protection of Vertebrate Animals Used for Experimental and other Scientific Purposes. *Council of Europe, European Treaty Series 123*. Strasbourg, 1986 (http://conventions.coe.int/treaty/en/Treaties/Word/123.doc).

Council of Europe (2004). Draft Appendix A of the European Convention for the Protection of Vertebrate Animals used for Experimental and other Scientific Purposes (ETS No. 123) Guidelines for accommodation and care of animals (Article 5 of the Convention) to be approved by the multilateral consultation. GT 123 (2004) 1. Final version for adoption. Strasbourg, 6 December 2004 (http://www.coe.int/T/E/Legal_affairs/Legal_co-operation/.

Dye, T.R. (2001). *Top Down Policymaking*. Chatham House Publishers, Inc. USA.

Höfer, T., Gerner, I., Gundert-Remy, U., Liebsch, M., Schulte, A., Spielmann, H., Vogel, R., Wettig, K. (2004). Animal testing and alternative approaches for the human health risk assessment under the proposed new European chemicals regulation. *Archives Toxicology 78,* 549-564.

Reinhardt, V. (2002). Comfortable Quarters for Primates in Research Institutions. In V. und A. Reinhardt (eds.), *Comfortable Quarters for Laboratory Animals*. 9th ed. Animal Welfare Institute PO Box 3650, Washington, DC 20007, 65-77. (http://www.awionline.org/pubs/cq02/cqindex.html).

Ruhdel, I. W. and Sauer, U. G. (1998). Primate experimentation. A report on the use, supply and housing conditions of primates used for scientific purposes within the European Union. European Coalition to End Animal Experiments, London. 80 pp.

Russell, W. M. S. and Burch, R. L. (1959). *The principles of humane experimental technique. Methuen*. London.

Schiffelers, M.-J., Hagelstein, G., Harreman, A., van der Spek, M. (2005). Regulatory Animal Testing – A Survey of the Factors Influencing the Use of Animal Testing to Meet Regulatory Requirements. Science Shop for Biology, Report No. P-UB-2005-10, Utrecht University, The Netherlands, 82 pages (http://www.bio.uu.nl/scienceshop/).

Correspondence to

Dr. med. vet. Ursula G. Sauer
Akademie für Tierschutz
Spechtstrasse 1
85579 Neubiberg
Germany
Tel.: +49 89 600 201 0
Fax.: +49 89 600 291 15
e-mail: ursula.sauer@tierschutzakademie.de

Workshop 3.2
Establishing the 3Rs principle around the world

Establishing the Three Rs Principle: A Plea for an International Severity Standard

Andrzej Elzanowski
Poland

Summary
Severity (harm, invasiveness) scales focus the attention of both regulators and researchers on highly invasive experiments and thus help to promote the 3Rs, especially the refinement. A severity scale is critical to the quality of ethical review processes (ERPs) in limiting arbitrariness and situational opportunism of assessments, defining the acceptability ranges for some animal uses, especially in education, and differentiating justification requirements. An international severity scale would improve ERPs in some countries and put pressure on others to institute them. An international severity scale would also facilitate the promulgation of other international requirements and restrictions that are applicable only within appropriate severity grades and that otherwise would be impractical, questionable or vague. I therefore advocate finding an appropriate international organisation to negotiate and propose an international severity standard and to launch an international campaign to promote its acceptance.

Keywords: animal welfare, international standards, animal research ethics, harm/pain/severity scale, refinement

Introduction

With many countries worldwide having little if any regulations of animal research and testing, any reasonable *minimum* standard adopted by an international committee is better than none, both in terms of its immediate effects when used by animal advocates, and as a draft for future improvements. The emphasis on the minimum is because any international standardisation initiative must not contravene more elaborate or restrictive standards already adopted by some progressive countries. However, in opposition to the widespread uncritical deference to religion and cultural diversity, the universal ethics of respecting all sentient beings may and should override any tradition of animal abuse.

There are at least five domains of standards that are jointly necessary to establish the 3Rs throughout the World: procedures (experimental and veterinary), housing and transport, ethical assessment and review, education and training, and enforcement. Only the two of them, the experimental and the housing and transport domains, have so far attracted considerable attention and effort. The only official world standards, for euthanasia and humane endpoints (International Council for Laboratory Animal Science, see www.iclas.org/) are in the procedural domain, and both are special cases insofar as both are concerned with death, which is an ultimately invasive procedure but can, if an appropriate care is taken, be administered quickly with only momentary or no suffering.

Animal research encompasses procedures ranging from non-invasive and prima facie benign procedures (such as handling, weighing, taking urine samples) to procedures inflicting extreme suffering and/or death. All animal research should be subject to ethical review and regulated, because even essentially harmless procedures can be harmful if improperly performed (Orlans, 1993), which is particularly true of wildlife research. On the other hand, one must not apply the same restrictions and requirements to, say nutritional experiments and neurosurgical procedures, because this would needlessly hamper science, antagonise research community, and ultimately backfire at all ethical oversight and legal control of animal research. In the

long run, over-regulation is known to be counterproductive for complex, intellectually demanding activities and this is certainly true of scientific research.

For that reason, among dozens of standards that are missing and necessary for the implementation of 3Rs, I advocate giving the top priority to the development of an international severity scale because it is badly needed as a yardstick for calibration of many other standards which turn out to be severity-sensitive. A severity scale is a conceptual tool for extending ethical oversight to all research on and with animals but without unduly hampering it. It is also needed to improve ethical review processes (ERPs) in some countries, especially in the US, and to have them started in others (especially in Russia and China).

A severity scale has another, psychological and thus evasive but potentially powerful educational effect on researchers if they have to assess and thus to confront themselves the harm they intend to do to their animal subjects. For example, the current Polish application forms require the researchers to determine the severity grade first, but a local ethics committee may correct the rating (in a special field). As a result, the researchers' rating is usually correct, which contrasts with the common practice of underrating the severity of one's own experiments (Orlans, 1993).

The importance of severity scales for ethical review processes

The current ideal for the ethical evaluation of animal research is best represented by Bateson's cube (Bateson, 1986) with its three variables, quality of research, medical benefit, and animal suffering, which all determine the ethical acceptability of a research project. It reflects the present western compromise between the majority of research community that conveniently embraces the anthropocentric dogma (Porter, 1992), and moral concerns of general public, supported and authenticated by the majority of lay philosophers. The ensuing compromise deserves the label of controlled speciecism: animals are harmed and killed in the name of substantial human benefits but at least animal suffering (and, to a various degree, life itself) should be reckoned with and not dispensed for trivial or questionable purposes. In reality, the control of speciecism varies between countries and animal species, and the Bateson's cube, which contains a space that is forbidden despite highest research quality and benefits, represents an ideal that may be approached only by a few nations and only with respect to primates, cats, and dogs.

In order to be operational, the Bateson's cube requires a scale or calibration for each of the three variables. Unfortunately, a precise quantitative scale does not seem possible for any of the three variables and the only way make them comparable is to agree on their gradation. It is highly desirable to have a graded classification for each of the variables (Porter, 1992) especially because an average human committee, with a majority of members with vested interests in continuing animal research, tends to overestimate human benefits and to underestimate non-human harms.

For the sake of implementing 3Rs, it is, therefore, essential to have the experimental procedures categorised by their severity grade prior to the evaluation of any invasive experiment. The gross underestimations of animal harms stem not only from largely unarticulated speciecism, but also from situational opportunism of assessments, with cronyism as a major but not the only factor. Situational opportunism stems from the pervasive net of personal, professional, and administrative interdependencies between the members of ethics committee as well as their ties with the outside colleagues who apply for approval of their projects. In Poland, we managed to keep, so far, our local ethics committees inter-institutional, with representatives from several independent research organisations, which reduces, but by far not eliminates, the impact of situational opportunism (which I can attest as a member of a local ethics committee).

While it is essential to make sure that even non-invasive handling of animals is done properly by responsible individuals, and thus to subject all animal research to ERP (which is the case in some countries including Poland), it would be counterproductive to require the same levels of justification for non-invasive or only slightly invasive procedures (such as blood sampling) as for heavy survival surgeries or acute toxicity testing. I therefore propose, as a rule for the ERPs, that requirements of project justification be severity-sensitive. This rule is implemented in the application forms established by Poland's National Ethics Committee, which require addressing clinical applications and documenting database searches only for two highest severity levels.

The use of live vertebrates in education is notoriously contentious, as exemplified by the admitted failure of the (really hard!) Working of Party of the Institute of Medical Ethics (UK) to reach consensus with regard to "invasive uses of animals" in undergraduate teaching (Smith and Boyd, 1991). This difficulty could be largely resolved if the use of animals could be limited to well-defined levels of severity although some disagreement as to the assessment of educational benefits will probably persist. Poland's first National Ethics Committee ruled in 1999 that only the lowest severity grade is acceptable for educational and training purposes and the rule seems to have been generally respected (except for a few objections from veterinary schools).

Severity grades as qualifiers of procedural standards

The harm scales permit to calibrate all restrictions and requirements to the harm to be inflicted by generic procedures, and thus to sensibly regulate all animal use in research, testing and education. Hence the enormous importance of a severity scale as a superstandard for some procedural standards that cannot be sensibly proposed without using severity grades as qualifiers. Without attempting a comprehensive review of all procedural standards that are missing for the effective implementation of 3Rs, the following examples suffice to demonstrate the significance of having an international severity scale as a superstandard for making research restrictions commensurate with the excepted harms.

Among many missing international standards are those defining the extent of veterinary supervision of experiments and the limits to repeated use of animal subjects. Both can be sensibly

proposed only with reference of severity grades inasmuch as requiring veterinary supervision for, or forbidding a repetition of some least invasive procedures would be impracticable and sometimes absurd.

Largely neglected are any regulations of wildlife research, although trapping and other more or less invasive experimental methods have been widely used for several decades (Cuthill, 1991) and the current widely used radio-tracking techniques (Mech and Barber, 2002) raise most serious concerns. The Canadian Council on Animal Care (CCAC) is the first major national agency that thoroughly addressed this issue (Griffin and Gauthier, 2004). Wildlife research regulations should be conceived as highly severity-sensitive because their enforcement in the field heavily depends of the levels of acceptance by the specific community of field naturalists and conservationists.

An outlook on national severity standards

Worldwide, a severity (harm, invasiveness) scale has by now been adopted by at least nine countries (tab. 1) including most Western countries leading in research.

Several different approaches to defining severity scales (tab.1) have been reviewed by Orlans (2000b) and Purves (2000). The latter mistook the Swiss standard for the retrospective assessment of severity (Bundesamt für Veterinärwesen, 1994b) for the German scale and listed the classification of animal experiments by Australia's National Health and Medical Research Council (NHMRC) among pain and distress classification systems although it cannot be regarded as such. The Australian classification is based on a mix of two criteria: kind of experimental procedure and some of its qualitative consequences for animal welfare. The Australian classification may be useful for devising a severity scale but cannot be used as such because it does not define any pain and distress grades or levels across various kinds of procedures, which is what a severity scale is about. It is regrettable that Australia did not adopt the excellent standard developed by the Australian and New Zealand Council for the Care of Animals in Research and Teaching (Mellor and Reid, 1994).

The majority of five countries, including New Zealand, Switzerland, Germany, Poland, and Sweden, use a ternary scale, that is, categorise all acceptable invasive procedures into three severity grades. The Netherlands used to have a ternary scale since 1979 (Orlans 1996); this scale is said to be currently revised. Two countries, Canada and UK, use a quaternary severity scale, that is, categorise all acceptable invasive procedures into four severity grades. In addition to 3 or 4 basic levels, some countries recognise a zero level for non-invasive or painless (which does not mean harmless) procedures and some countries identify extreme and/or unacceptable procedures in a special category (X in tab. 1).

The categorisation of animal experiments used by the US Department of Agriculture's Animal and Plant Health Inspection Service (APHIS) was designed in the 1970's (Orlans 1993, 2000b) as a catch-all formula for reporting rather than ethical review purposes. It is not really a severity scale (hence it is not listed in tab. 1) as it is based on the use or non-use of anaesthesia (as deemed appropriate for various reasons) rather than severity *per se*. Three categories of experiments are defined in the column headings in the Annual Report of Research Facility form (USDA/APHIS, 1991): experiments involving no pain, distress or (sic) use of pain-relieving drugs (column C), experiments involving pain or distress which are alleviated by anaesthetics, analgesics or tranquilizers (column D), and experiments for which administering such drugs for relief "would have adversely affected the procedures, results or interpretation of the teaching, research, experiments, surgery or tests" (column

Tab. 1: Severity scales worldwide. 0 denotes non-invasive procedures on living animals. X denotes extreme or unacceptable procedures if recognised as such in the cited national standard.

Country	Severity grades	Since	Source
The Netherlands	under revision[1]	1979	Orlans (1996)
Finland	2	1985	Laboratory Animal Act (1076/1985)
Canada	4[2]	1987	CCAC (1991)
New Zealand	0+3+X[3]	1988	Mellor and Reid (1994)
Switzerland	0+3	1994	Bundesamt für Veterinärwesen (1994a)
Poland	0+3+X	2000	Krajowa Komisja Etyczna (1999)
UK	4+X	2000	Home Office (2000
Germany	0+3[4]	2000	Bundesministerium der Justiz (2000)
Sweden	3[5]	2001	Djur Skydds Myndigheten (2001)

[1] The Netherlands used to have a three-level severity scale.
[2] Of the four counted levels (B through E), the lowest (B) includes some procedures categorized elsewhere as 0 and the highest (E) includes some procedures categorized elsewhere as X.
[3] Applied in each of the five harm ("welfare compromise") domains (Mellor & Reid, 1994)
[4] Used in the application form but not explicitly defined; duration of a procedure is considered separate from the severity, as the form requires its specification (1 day, 1-7 days, 7-30 days, and above 30 days) in addition to a severity grade.
[5] Sweden used a six-level scale in the years 1979-1989.

E). It is only for the latter (column E) category that an approval by an Institutional Animal Care and Use Committee (IACUC) is required. The requirements stated in the Annual Report form are interpreted and elaborated in the APHIS policy, the most recent version of which (USDA/APHIS, 1997) defines a painful procedure as any procedure that can reasonably be expected to cause more than slight or momentary pain". In fact, only the most invasive procedures such as terminal surgery, Freund's complete adjuvant, ocular and skin irritancy testing, extensive irradiation, long-term food or water deprivation, noxious electrical shock, paralysis or immobility are considered to be painful. The majority of experiments is liberally categorised as causing only slight or momentary pain, lumped with non-invasive procedures, and thus effectively ignored. This regulation may have been progressive in the 1970's, i.e., at the time of general callousness toward animal suffering in the US and most other countries, but today it is outdated and inappropriate as a severity standard for use in ERP. It is most deplorable that the US research community has contravened sincere efforts on the part of its progressive minority (as well as USDA in 1987) to establish an adequate severity standard (Orlans, 2000b), which puts the US behind many other Western nations in a key ethical matter.

A two-level severity scale has also been established in Finland by its 1076/1985 Laboratory Animal Law, each level ("class") being defined by several examples in a regulatory act (Decision 477/1986) issued by Finland's Ministry of Agriculture and Forestry, and a grade 0 is also used in the ERP for experiments that "do not cause harm or discomfort" and thus do not require any permits (Kai Pelkonen, pers. comm.).

There are countries in which no severity scale is prescribed by any law or regulation and yet some severity assessment is done in the ERP. For example in Singapore the veterinarian in charge of the care of experimental animals is expected to assess the severity of the projected experiment (Su Hua Leow, pers. comm.) although the elaborate Guidelines on the Care and Use of Animals for Scientific Purposes, as instituted by the National Advisory Committee for Laboratory Animal Research in 2004, do not define a severity scale.

Ternary scales

The most refined ternary scheme has been devised by ANZCCART to consider pain and distress from all sources which are conceptualised as five domains (Mellor and Reid, 1994) and then to add them in calculating the summary severity grade. This scheme should be used together with the Swiss rating of procedures (see below) by all regulatory bodies and standard setting panels for assigning examples to each level, but may be too complex for an average local ethics committee.

The Swiss Federal Veterinary Office (Bundesamt für Veterinärwesen, 1994a) distinguishes four expected (pre-experimental) severity grades starting with grade 0 for non-invasive procedures (such as ethological observations). Examples for each grade are provided in 12 areas: restrictions of movement and diet, interference with reproduction and development, sampling and surgery, exposure to physical agents, pharmacology and toxicology, microbiology and parasitology, immunology, analgesia and inflammation, heart and circulation disorders, hormonal and metabolic disorders, neurological and behavioural disorders, and cancer. A separate standard (Bundesamt für Veterinärwesen, 1994b) defines procedures and symptoms for common laboratory species for the retrospective determination of the actual severity grades in ongoing projects. Retrospective review is essential for an effective oversight and has recently been addressed in the European Commission and Britain (LASA, 2004). Unfortunately, the concept is completely unknown in Poland (and probably most other new EU member states) and adds to the long list of missing standards. Retrospective review would also benefit from a severity scale as projects involving the most invasive experiment should obviously be revisited first. This is because ethics committees in many countries may not have enough capacities to conduct retrospective review of all approved projects.

Poland's National Ethics Committee for Animal Experimentation (Krajowa Komisja Etyczna do Spraw Doświadczeń na Zwierzętach), empowered by the 1997 Animal Welfare Act to set up standards for the ERP, adopted in 1999 an essentially ternary scheme (tab. 2) which is based on the scale once proposed for the US by the Scientists' Center for Animal Welfare (Orlans, 1990, see also Fraser, 1990). From the very beginning the scheme has been effectively enforced through application forms, which require the severity grade to be specified by the investigator, and make several requirements dependent upon the severity grade, e.g., only the lowest severity grade is permissible for educational experiments, and clinical implications need to be specified for higher grades. The use of severity grades has been mandated by Poland's new Animal Experimentation Act of January 21, 2005 (Dziennik Ustaw 05.33.289 of 24 February 2005) which requires (in Art. 20) a determination of the severity grade in the permit application and delegates (in Art. 28) the definition of the severity scale to the National Ethics Committee, the latter provision making it inevitably dependent on the political vicissitudes of this body.

Tab. 2: Severity scale instituted by Poland's National Ethics Committee (Krajowa Komisja Etyczna, 1999) and largely based on the scale once proposed by the Scientists' Center for Animal Welfare (Orlans, 1990).

Grade 1	Noninvasive
Grade 2	Light, momentary pain/distress or minor discomfort for a long duration.
Grade 3	Moderate pain or distress.
Grade 4	Acute pain or distress or irreversible physical or psychological impairment.
Grade X	Unacceptable extreme suffering.

The ceding post-communist government has just appointed a new, egregiously inappropriate National Ethics Committee, with an immediate effect of the Scale of Invasiveness having been removed from the Web page of the Ministry of Science.

After the unfortunate abandoning its once pioneering severity scale (Orlans, 2000b), Sweden has now reinstated a simple three-grade scale, with each grade defined by a dozen of known examples by the Swedish Animal Welfare Agency (Djur Skydds Myndigheten, 2001).

Three severity grades are used in Germany's application form for the animal experimentation permit (to a state or *Land* government), which is part of the ordinance issued the Federal Ministry of Justice (Bundesministerium für Justiz, 2000) pursuant to the 1998 update of the German Animal Welfare Act. However, there is no explicit definition of these grades, which are meant to scale only the intensity but not the duration of procedures. The latter has to be specified separately in the same form in terms of days.

The Japanese "Standards relating to the care and management etc. of experimental animals" (issued in 1980 by the Prime Minister's Office) do not contain any provisions for the use of severity scales. However, some institutions voluntarily use the ternary scale devised in the eighties by the US Scientists' Center for Animal Welfare (Yukihisa Matsuda, pers. comm.).

Quaternary scales

The Canadian Council on Animal Care defines five categories of invasiveness (CCAC, 1991). However, the lowest of them refers to those invertebrates that are considered to be non-sentient and to tissue cultures, and thus will be ignored for the purpose of comparisons to other national standards. The lowest of the remaining four grades (B) comprises a mix of procedures that are *per se* non-invasive, such as temporary maintenance of a herd under a production management system to be tested, and clearly invasive procedures such as blood sampling, simple injections or approved methods of euthanasia involving rapid unconsciousness. The label "minor stress or pain of short duration" for the following level C underestimates the severity of at least some of the listed procedures (such as "stressful restraint"). The next level D is again mixed as admitted by its title "moderate to severe distress or discomfort". The top level E includes several procedures that are widely considered unacceptable (such as the use of muscle relaxants and paralytic drugs without anaesthesia or burn infliction on non-anaesthetised animals) as well euthanasia methods not approved by the CCAC. And yet the entire category is not identified as unacceptable (and thus is not identified as grade X in tab. 1). Because of the mixed contents of its categories, the Canadian scale lacks transparency and tends to underestimate, at least verbally, the severity of listed procedures.

In the UK, the Home Office (2000) specifies the four categories in its *Guidance to the Operation of the Act*. The lowest is called "unclassified" and includes terminal experiments entirely under general anaesthesia, and the remaining three are called mild, moderate, and substantial. In addition, "the Secretary of State will not license any procedure likely to cause severe pain or distress that cannot be alleviated [Section 10(2A)].", which is what is categorised in table 1 as forbidden or extreme procedures (grade X). Thus the British scale seems to be best conceptualised as 4+X (tab. 1) even though classifying terminal anaesthesia in a category of its own seems somewhat inconsistent because there are many procedures that are comparable in severity to anaesthesia, and euthanasia *per se* is not covered by the British scale. This scale is used for an *upper limit* on the severity acceptable for an individual experimental protocol. In addition, there are severity bands for programmes of work where they indicate the *average severity*, taking into account numbers of animals affected, intensity and duration of the adverse effects, for all animals expected to be used during the whole programme.

An evaluative outlook

The great advantage of ternary schemes is that they are simple and nearly self-explanatory: a procedure that is not comparable to either the most or the least invasive examples falls in the middle. Its simplicity also makes it independent of frequent updates and elaborations. There are inevitable complaints about the relative imprecision of a ternary assessment (U. Hansson, pers. comm.), especially in Sweden, where a refined six-level scale was formerly used. However, the refinement of laws should not go too far beyond the possibilities of their enforcement. Therefore, simple ternary scales (similar to the Polish and the current Swedish scale) seem to be the best choice for countries such as new EU member states, with little experience and resources to carry out ERP. A two-level scale, with all experiments categorised as either mild or severe, seems too simplistic as many experiments that cause considerable suffering may fall in the mild(er) category, although a judicious use a two-level scale is probably much better than none.

Whatever the number of severity grades, it is of great importance to have a category for procedures that are harmless by laboratory standards as even the least invasive procedures may turn harmful if improperly executed (as in the case temporary captivity followed by a release at the wrong time or place). This is especially true for wildlife research, which is usually done in the name of nature conservation and thus looks animal-friendly or at least benign to the public. In fact, research on wild animals frequently leads to slow and painful death, e.g., when caused by a too tight telemetric collar, which is on the average worse than death from most natural causes.

A plea to speed up international standardisation of severity scales

In view of the urgent needs and considerable number of good regional standards, the process of turning them into World standards needs to be substantially accelerated. Each of the five categories of needed standards requires different expertise and philosophical perspective, calling for involving more organisations in the process, but there may be not enough of them to do the job. For example, while the ICLAS and other laboratory animal associations are well prepared to propose experimental standards and the housing and transport standards, the three other domains should be tackled by bodies representing broader

expertise including appropriate psychological and philosophical perspectives. One of several candidate organisations (Purves, 2000) is the International Organization for Standardization (ISO) which is currently embarking on social responsibility issues (Frost, R., 2005. ISO Press Release 947: http://www.iso./oso/en/commcentre/pressreleases/) and thus could be approached to support the development of a standard severity scale as a good exercise in applying public responsibility in science and biotechnology.

References

Bateson, P. (1986). When to experiment on animals. *New Scientist 109*, 30-32.

Bundesamt für Veterinärwesen (1994a). Einteilung von Tierversuchen nach Schweregraden vor Versuchsbeginn (Belastungskategorien). Allgemeine Leitsätze und Beispiele zur analogen Klassierung weiterer Versuche. Bundesamt für Veterinärwesen, Information Tierschutz 1.04.

Bundesamt für Veterinärwesen (1994b). Retrospektive Einteilung von Tierversuchen nach Schweregraden (Belastungskategorien). Bundesamt für Veterinärwesen, Information Tierschutz 1.05.

Bundesministerium der Justiz (2000). Allgemeine Verwaltungsvorschrift zur Durchführung des Tierschutzgesetzes. *Bundesanzeiger 52* (36a): 1-16.

CCAC (1991). Position Statement: Categories of Invasiveness in Animal Experiments. Ottawa, Canadian Council on Animal Care. (Reprinted in Olfert, E. D., Cross, B. M., and McWilliam A. A., eds. *Guide to the Care and Use of Experimental Animals Vol. 1*, 2nd ed., 1993).

Cuthill, I. (1991). Field experiments in animal behaviour: methods and ethics. *Animal Behaviour 42*, 1007-1014.

Djur Skydds Myndigheten (2001). Etisk prövening av djurförsök. Vägledning till ansökan.

Fraser, A. F. (1990). An analysis of suffering. In B. E. Rollin & M. L. Kessel (eds.), *The Experimental Animal in Biomedical Research*, Vol. 1. Boca Raton, CRC Press.

Griffin, G. and Gauthier, C. (2004). Incorporation of the principles of the Three Rs in wildlife research. *ATLA 32 (Suppl. 1)*, 215-219.

Home Office (2000). *Guidance on the Operation of the Animals (Scientific Procedures) Act 1986*. London, HM Stationary Office.

Krajowa Komisja Etyczna (2000). Skala Inwazyjności Badań na Żywych Kręgowcach. Komitet Badań Naukowych (www.kbn.gov.pl).

LASA (2004). *Guidance notes on retrospective review: a discussion document prepared by the LASA Ethics and Training Group*. M. Jennings and & B. Howard (eds.) Laboratory Animal Science Association.

Mech, L.D. and Barber S. M. 2002. A critique of wildlife radio-tracking and its use in national parks. A report to the U.S. National Park Service.

Mellor, D. J., Reid, C. S.W. (1994) Concepts of animal well-being and predicting the impact of procedures on experimental animals. In Improving the Well-being of Animals in the Research Environment. Australian and New Zealand Council for the Care of Animals in Research and Teaching (ANZCCART), Glen Osmond, South Australia.

Orlans, B. F. 1990. Animal pain scales and public policy. *ATLA 18*, 41-50.

Orlans, B. F. 1993. *In the Name of Science / Issues in Responsible Experimentation*. Oxford University Press.

Orlans, B. F. 1996. Invasiveness scales for animal pain and distress. *Lab Animal 25*, 23-25.

Orlans, B. F. 2000a. Practical applications of animal harm scales: international perspectives. In M. Balls, A-M. van Zeller and M. E. Halder (eds.), *Progress in the Reduction, Refinement and Replacement of Animal Experimentation*. Elsevier Science.

Orlans, B. F. 2000b. Public policies on assessing and reporting degrees of animal harm: international perspectives. In M. Balls, A-M. van Zeller and M. E. Halder (eds.), *Progress in the Reduction, Refinement and Replacement of Animal Experimentation*. Elsevier Science.

Porter, D. G. 1992. Ethical scores for animal experiments. *Nature 356*, 101-102.

Purves, K. E. 2000. Global harmonization of pain and distress classification systems: current analysis and opportunity for change. *Journal of Applied Animal Welfare Science 3*, 49-61.

Smith J. A. and Boyd K. M. (eds.) 1991. *Lives in the Balance / The Ethics of Using Animals in Biomedical Research. The Report of a Working Party of the Institute of Medical Ethics.* Oxford University Press.

USDA/APHIS 1991. Annual report of research facility. APHIS form 7023.

USDA/APHIS 1997. Animal Care Policies (Painful Procedures – Policy # 11).

Acknowledgments

I thank Prof. Horst Spielmann (ZEBET, Berlin) for his generous invitation to the Congress and his great help in promoting 3Rs in Poland. I greatly appreciate the input from Robert Baker (Department of Primary Industries and Resources of South Australia), Derek Fry (Home Office, UK), Ulrika Hansson (Stiftelsen Forskning utan Djurförsök / The Swedish Fund for Research Without Animal Experiments), Roman Kolar (Akademie für Tierschutz, Deutscher Tierschutzbund), Su Hua Leow (Agri-Food & Veterinary Authority, Singapore), Yukihisa Matsuda (Akita University), Kai Pelkonen (Ministry of Agriculture and Forestry, Finland), and Dan Scholfield (Animal Welfare Information Center, USDA).

Correspondence to

Andrzej Elzanowski
Department of Zoology
University of Wroclaw
21 Sienkiewicz Street
50335 Wroclaw
Poland
e-mail: elzanowski@biol.uni.wroc.pl

Current Status of Establishing the 3Rs Concept in India

Shiranee Pereira
CIBA, Indian Council of Agricultural Research, Chennai, India

Summary
The CPCSEA (Committee for the Purpose of Control & Supervision of Experiments on Animals) is a statutory body of the Government of India and draws its powers from the Prevention of Cruelty to Animals (PCA) Act of 1960. The Act states that the "duty of the committee is to take all such measures as may be necessary to ensure that animals are not subject to unnecessary pain and suffering before, during and after the experiment" and "experiments on animals be avoided wherever it is possible to do so if other devices and the like may equally suffice". This can be traced as the first legal indication of the concept of alternatives in India. The PCA Act of 1960 was conceptualised by Rukmini Devi Arundale, theophist and philosopher, a contemporary of the founders of the classical concept of the 3Rs – W. M. S. Russel and R. L. Burch. The "Breeding of and Experiments on Animals (Control & Supervision) Rules 1998" were notified in the gazette of India in 1998. These rules further reiterate the concept of alternatives. The CPCSEA proactively works to propagate the concept of the 3Rs and an independent expert committee has been constituted to suggest and implement alternatives in research, education and drug testing. There is a profound difference today, especially in the use of animals in education, in the use of equines in the production of immunobiologicals, in the manufacture of tissue culture vaccines, LD_{50}, etc. The CPCSEA, in 2004, has further officially accepted the concept of the 4th R –"Rehabilitation" of laboratory animals, stating that investigators have a moral responsibility to take care of laboratory animals after use.

Keywords: India, PCA Act, CPCSEA, three Rs, the 4th R

Introduction

In India, the concept of alternatives, though not rigidly referred to as the concept of the 3Rs, had its first indications in the Prevention of Cruelty to Animals (PCA) Act of 1960. Undoubtedly borne from India's religious and cultural traditions and philosophy of "Ahimsa" or "non-violence", the reference to the use of alternatives to animals in education and research was also a reflection of the personal beliefs and philosophy of Rukmini Devi Arundale, theophist and philosopher, the architect of the PCA Act.

Most Indians regard all forms of life as equally important, considering them incarnations of a single energy or life force. On a still higher plane, Hindu mythology endowed the gods with animal attributes. The lord Ganesh (the elephant god) is inseparable from his little mouse friend, Shiva cannot be worshipped without his garland of serpents, Krishna cannot be pictured without his cows, Goddess Saraswati, the goddess of wealth, is inseparable from her swan, and Lord Rama could not have been Rama without the Hanuman, the monkey god.

Hence, in a country where animals and spirituality are so intermeshed, it was a paradox and shock when the CPCSEA (Committee for the Purpose of Control & Supervision of Experiments on Animals), the statutory body of the government of India, under the chairpersonship of Madam Maneka Gandhi, in the year 2000, exposed the appalling conditions of animals and animal houses in laboratories across the country. Unfortunately, laboratories in independent India were referred to as "temples" and scientists as the "high priests". But what transpired in these temples was far from god or religion. The sacredness of life and the concepts of Dharma (Mercy), Karma (as you sow, you shall reap) and Ahimsa (non-violence) had no place here. It was a scenario in which thousands of animals were bled, bludgeoned, maimed and mutilated, starved and thirsty, caged and concussed, for no purpose or panacea, and their deaths saw no claims, patents, cures or curbs (Veeraraghavan, 2004). To quote Maneka Gandhi, Ex-Chairperson of the CPCSEA, "India needs to look at its scientists and put a stop to the present state of directionless, unmeaningful, sub-standard and cruel research. How valid is biomedical research in India? This is not a human vs. animals issue. The issues that I have raised having nothing to do with the cruelty to animals by itself. They are about the lack of science." (Pereira, 2002)

The Evolution, history and current status of the concept of 3Rs in India

CPCSEA was founded by an act of the Indian parliament and draws its powers from the Prevention of Cruelty to Animals (PCA) Act of 1960. Section 17 (1) of the Act states that the "duty of the committee is to take all such measures as may be necessary to ensure that animals are not subject to unnecessary pain and suffering before, during and after the experiment", which in spirit is the concept of the 3Rs of Russel and Burch (1959). The Act further indicates the replacement of animals in experimentation stating that "Experiments on animals be avoided wherever it is possible to do so if other devices and the like may equally suffice".

For over three decades after the first CPCSEA was constituted in 1964, little was done to alleviate the suffering of laboratory

animals in India. Apparently, the committees failed to impose the laws and a dismal scenario continued in laboratories across the country (Pereira et al., 2004).

In 1998, under the committed chairpersonship of Maneka Gandhi, a pro-active secretariat of the CPCSEA was created. With the power to promulgate its own laws to ensure the humane and ethical use of animals in research and education, the CPCSEA in 1998 notified the "Breeding of and Experiments on Animals (Control & Supervision) Rules 1998", which were further amended in February 2001.

From 1999, CPCSEA enforced the laws and for the first time in the history of independent India there is now a profound difference, especially in the use of animals in education, in the use of equines in the production of immunobiologicals, in the manufacture of tissue culture vaccines, LD_{50}, etc. The CPCSEA, in 2004, further officially accepted the concept of the 4th R – "Rehabilitation" of laboratory animals, stating that investigators have a moral responsibility to take care of laboratory animals after use.

Education

In recent years there has been an awakening and a realisation in academia of the profound negative impact and futility of using animals in education from both a pedagogic and a psychological point of view. It has been recognised that fear, anxiety and revulsion are imposed on the psyche of a young mind by compulsory vivisection. The questions posed were: "Is education legitimising harm and death? To a young mind, is the biology class the first lesson of violence? Can the beauty of life and living things be taught through death? Is it not a paradox of education?"

An issue bordering on human rights and a child's right or desire not to kill or harm, however, needs more compulsive thought and urgent attention by further education boards in India.

In August 2000, the Indian School Certificate (ISC) Board for Education, one of the largest national bodies for secondary education in India, decided, "Dissection of animals will not be a component of biology practicals from the year 2000".

In April 2001, the Central Board of Secondary Education (CBSE), India's largest national body for primary and secondary education decided to "delete all experiments relating to dissection of animals in biology practicals in the senior school curriculum" in all CBSE schools across India.

In 2003, the Pharmacy Council of India, the academic body that governs the 300 odd pharmacy colleges in India, issued a directive to adopt the use of the "Expharm CD", which is a full and direct replacement to the use of animals in the teaching of pharmacology, on the request of the CPCSEA.

In August-September 2004, a training course on the use of alternatives in education was organised in 11 cities in India. Over 800 academia across India, veterinary medicine, pharmacology, life sciences and medical universities, were invited to send their representatives to a training workshop and seminar entitled "Alternatives, Animal Welfare and the Curriculum". Over 400 teachers were trained in the use of about 22 alternatives, which can directly replace the use of animals in veterinary medicine, pharmacology, medical and life science education.

In 2005, as a result of its long-term liaison with the CPCSEA, the Department of Life Sciences of Bharathidasan University, one of the premier universities in India, introduced a 95% ban on the use of animals in life science teaching.

Research

With the establishment of an active secretariat of the CPCSEA, the "Breeding of and Experiments on Animals (Control & Supervision) Rules 1998" was imposed on every establishment and breeder using or breeding animals for the purpose of education or research. The Concept of the 3Rs was introduced as the inevitable scientific solution to credible research. In the year 2000, the CPCSEA put together a national subcommittee for the promotion and propagation of the concept of alternatives in education, basic biomedical research and regulatory testing. This resulted in over 900 laboratories registering with CPCSEA. Ethics committees have been constituted in all registered institutes, honorary CPCSEA nominees have been appointed in these institutes, a special expert committee at the national level has been appointed to scrutinise and approve the use of large animals, a national "Good Laboratory Practice" document was introduced, a mandatory protocol for the care and use of equines in the production of immunobiologicals has been imposed, and the credo of 3Rs was introduced as part of every project proposal and as a legal requirement. Validated alternatives (ECVAM and ICCVAM) have been recommended to regulatory authorities, the Ministry of Health has been urged to place a national ban on the Semple vaccine, and hundreds of laboratory animals have been rehabilitated.

In March 2004, a representation from the scientific community to the CPCSEA requested that existing national norms on laboratory animal care and use be amended, so as to exclude rats, mice, birds and farm animals from all/any kind of regulation. Apropos this, the CPCSEA constituted a Consultative Group, comprising representatives from the Ministry of Health, Ministry of Science and Technology, National Research Councils, premier research institutes, philosophers and animal welfare personnel to consider this issue. The consultative group elucidated, first, the principles that should form the framework in order to review the existing norms and for the promulgation of new norms.

To facilitate the discussions, an order of relative sentience of the different species was first formulated. It was agreed on consensus that any being higher than a cockroach, representing the invertebrate, would require regulation. Thus mice, rats, birds and farm animals will *not* be excluded from regulation (Pereira et al., 2005). A further five new principles in elaboration of the 3Rs were formulated for the reduction of the use of animals, refinement by way of mandatory use of analgesics and anaesthetics, guiding principles if animals have to be euthanised and the mandatory need to rehabilitate laboratory animals after use.

The establishment of the concept of the 4th R – rehabilitation of laboratory animals

The concept of the 4th R - rehabilitation of laboratory animals was borne in India out of the urgent need to provide relief and succour to ailing animals in laboratories. Rehabilitation is undertaken when the need arises with the sole intention of alleviating any form of suffering or pain and/or to save the life of the animal (Pereira and Tettamanti, 2005). In 2004, the CPCSEA incorporated the concept of the 4th R into its official guidelines, stating that:

"Personnel using experimental animals have a moral responsibility for the animals after their use. Investigators are responsible for the aftercare and/or rehabilitation of animals post-experimentation, and may be permitted to euthanise animals only in the following situations:

- When the animal is paralysed and is not able to perform its natural functions, it becomes incapable of independent locomotion and/or can no longer perceive the environment in an intelligible manner.
- During the course of the experimental procedure, the animal has been left with a severe recurring pain, wherein the animal exhibits obvious signs of long-term extreme pain and distress.
- In situations where non-termination of the animal experimented upon would be life threatening to human beings or other animals.

Investigators are responsible for animals even after termination of the experiment and they have a moral obligation to ensure that experimental animals should be rehabilitated at the end of the experiment.

Costs of aftercare and/or rehabilitation of animals post-experimentation are to be part of research costs and should be scaled per animal in positive correlation with the level of sentience of the animals." (Anon, 2004)

In 2004, I-CARE (International Centre for Alternatives in Research and Education), an NGO working to promote the cause of the 3Rs and dedicated to the 4th R, was created to give positive reinforcement and presence to the use of alternatives in research and education in India.

It was created recognising the need of the hour and the fact that compulsion cannot sustain positive change. Sustained change can only be achieved if every individual of the scientific community is convinced that the use of alternatives is better and progressive, and that the humane use of animals in experimentation is innate to credible research. Teachers should realise that the use of alternatives in education is pedagogically superior, environmentally friendly and preserves the sensitivity of the student. The rationale behind the establishment of this centre is to empower individuals to make this change and to take positive initiative in making humane and responsible choices.

References

Anon (2004). Report of the Consultative Group on Review of the Norms and Practices for Regulation of Animal Experimentation. Ministry of Environment and Forests, Animal Welfare Division, Government of India.

Russel, W. M. S. and Burch, R. L. (1959). *The Principles of Humane Experimental Technique*. London: Methuen.

Pereira, Shiranee (2002). Labs or Torture hovels. *Manushi 132*, 12-16.

Pereira, Shiranee, Veeraraghavan, Sonya Ghosh, Sonya and Gandhi, Maneka (2004). Animal Experimentation and Ethics In India: The CPCSEA Makes a Difference. *ATLA 32*, *Supplement 1*, 411-415.

Pereira, Shiranee and Tettamanti, Massimo (2005). Ahimsa and Alternatives – the concept of the 4th R. The CPCSEA in India. *ALTEX, 22 (1)*, 3-6.

Pereira, Shiranee, Tettamanti, Massimo and Bodnariu. Alina (2005). "*Sentience – The basis for new laboratory animal welfare policies in India*." Poster presented at the International Conference "From Darwin to Dawkins: the science and implications of animal sentience". London, March 2005. Organized by The CIWF, UK.

Veeraraghavan, Prema (2004). Ahimsa Annihilated. AAVS Magazine, *Fall 2004*, 14-15.

Correspondence to

Shiranee Pereira
Consultant, CPCSEA, Animal Welfare Division
Ministry of Environment and Forests, Govt. of India
CIBA, Indian Council of Agricultural Research
75 Santhome Hg. Road
Chennai 600 028, India
e-mail: shiraneep@hotmail.com

The 3Rs in Brazil

Ekaterina A. B. Rivera
Biotério Central-Universidade Federal de Goiás, Goiânia, Brazil

Summary
In Brazil scientists, teachers, students and technicians have a new approach when working with animals due to the principle of the 3Rs. It is possible to clearly notice the changes in their attitudes towards animals, showing respect and treating them as sensitive beings and not as objects as they used to do some years ago.

Keywords: Brazil, three Rs

Introduction

Brazil is a young and continental country and for this latter reason you can find different economical, social, cultural and educational statuses. This principle also applies to the scientific field, where you have areas of excellence in some regions and in others very little knowledge of the same subjects. In relation to the 3Rs we had been working restlessly trying to reach the whole country. It is not an easy task because there are many different ways of thinking among users of animals.

Methods

A strategic plan, trying to spread the principle of the 3Rs, has been developed and it involves 4 different main areas:
1. Universities and Research Institutes
2. Brazilian Regulatory Agencies
3. Financial Agencies
4. Scientific Magazines and Journals/Congresses

1. Universities and Research Institutes – The first step was to try to achieve our goals through education and there is no better field to sow than at Teaching Institutions. The importance of the 3Rs in research, education and tests has been emphasised through lectures, workshops and seminars. Courses for post graduation students as well as for technicians are given in many Institutions. Various alternatives to the use of animals in teaching and research have been introduced in our Universities and Research Centers. The creation of Ethical Committees in almost every University has helped to improve the implementation of the 3Rs.
2. Toxicity testings in animals for the assessment of new chemicals are mandatory before their introduction in the country and are required by the Ministry of Health and the Ministry of Environment. These tests follow OECD and EPA guidelines.
3. The Brazilian Research Council and other Financial Agencies do not approve projects which involve animals if these projects are not submitted to an Ethical Committee. The same applies to some of our Scientific Magazines and Journals.

Conclusions

There has been a great change among users of animals and the culture of care is now a common practice in the great majority of our Universities and Colleges. This is due mainly to the implementation of Ethical Committees in almost every institution where animals are used. A sharp decrease in the number of animals used in teaching programs has been noticed and an increase of more than 75% in the use of alternatives instead of animals in practical classes. The students are aware that alternatives are the future and they prefer them instead of using animals, saying that they feel more comfortable not using animals in their classes.

The Brazilian Regulatory Agencies have a modern and very good legislation regulating the required tests to assess the risk of new chemicals. From time to time these regulations have been reviewed and there is a tendency to eliminate some of the tests in order to reduce the number of animals used. When the law was first issued in 1990, 12 tests were required; in 1994 and in 1996 some of the tests were only conditionally required or bibliography could be accepted (tab. 1). Our Regulatory Agencies are very interested in applying the 3Rs. They require the contract testing laboratories to work in a GLP system and to have well trained technicians. The purpose is to use less animals, looking for alternatives to traditional tests without compromising animal and human health.

Financial Agencies, mainly the Brazilian Research Council, and scientific magazines and journals have an important role in the improvement of the 3Rs. They require that scientists submit their papers to Ethical Committees before giving their projects or papers. Ethical Committees require the 3Rs to be observed otherwise the paper will not be published or the project will not get financial support.

In 1992 we promoted an International Congress of the Brazilian Laboratory Animal Association, and the main theme of

Tab. 1: Required tests with animals/Ministry of Health-Ministry of Environment

Test	14/03/1990	21/12/1994	15/101996
AcuteOral toxiciy	X	X	X
Acute Dermal	X	X	X
Acute inhalation	X	X	CR
Acute dermal irr	X	X	CR
Acute eye irrit	X	X	CR
Repeated 28 days in rodents			X
90 days oral rodents	X	X	CR
90 days oral dogs	X	X	CR/B
Chronic Toxicity*	B	B	
Metabolism			X
Sub-chronic inhalation	X		
Sub-chronic dermal	X	X	

*Decretos do IBAMA- 1990, 1994, 1996, Brazil

the scientific program was the 3Rs. During the Congress we also gave a prize for the best paper on alternatives in education and another one to alternatives in research.

Other seminars, congresses and workshops have the 3Rs as their main subject or at least they are part of the program.

Implementing the 3Rs has been continuous and hard work, but it has been worth doing it. The results are more and more encouraging keeping us stimulated to continue this work. Every day more and more people get in touch with us who are interested in the subject and asking how they can apply the 3Rs. We can say, with no doubt, that Brazil is a country which, in a near future, will be using less animals in research, testing and teaching, refining the procedures used and looking for alternatives to the use of animals whenever possible.

Thus, as mentioned before, Brazil will speak the same scientific language as other countries.

This is a descriptive work, based on data of some inquiries made by the Brazilian Veterinary Association and some personal observations and resports sent by people working in the area.

Correspondence to
Ekaterina A. B. Rivera
Biotério Central-Universidade Federal de Goiás
Goiânia, Brazil
e-mail: e.rivera@uol.com.br

Session 3.4
Policy implementation

Three Years of Animal Welfare in the German Constitution – the Balance from an Animal Welfare Perspective

Roman Kolar
Animal Welfare Academy, German Animal Welfare Federation, Neubiberg, Germany

Summary
The inclusion of animal welfare into the German Constitution in 2002 gained worldwide attention. Animal welfare organisations had been striving to reach this goal for more than a decade, and they had good reasons to do so: in several court cases the regulations of the German Animal Welfare Act had been overruled by basic rights laid down in the Constitution, such as the freedom of research and the freedom of education.
To review the situation, case studies focussing on court cases and assessments by members of ethics committees are used to demonstrate what consequences the change of the German Constitution has brought about.

Keywords: animal protection, animal welfare legislation, ethics committees

Introduction

Before the year 2002, the appropriate regulation of animal experimentation in Germany was severely compromised by the fact that the German Constitution did not include animal welfare as a specific objective, whereas freedom of (*inter alia*) arts, religion, education and research were, and still are, guaranteed as basic rights (Caspar und Schröter, 2003).

For that reason, High Court Decisions directly interfered with regulations laid down by the German Animal Welfare Act (Bundestag, 1998), as well as EU Council Directive 86/609 (Anon., 1986a), and the Council of Europe Convention ETS 123 for the Protection of Animals Used for Experimental and other Scientific Purposes (Anon., 1986b) in those cases where there was a conflict between animal welfare and constitutional rights.

Court cases before 2002

Freedom of education overrules the Animal Welfare Act
In 1993 a professor at the University of Giessen filed a lawsuit against the local competent authority's decision to not grant licenses for experiments on live rats for educational purposes (demonstration of glucose uptake) that he had applied for. In its decision on this case, the competent Supreme Administrative Court (*Verwaltungsgerichtshof*) of the City of Kassel came to the conclusion that the experiments were legal, even though it had been demonstrated that an alternative (a film showing the experiment) existed. It stated that animal welfare, based on ethical grounds, did not have constitutional status and therefore could not compromise the university teacher's basic right to freedom of education.[1]

German Federal Constitutional Court restricts ethical examination of animal research
In 1994, a professor at the University of Berlin filed a lawsuit against the local competent authority's decision to no longer

[1] Verwaltungsgerichtshof (VGH) Hessen, Kassel, Urteil vom 29.12.1993, AZ 11 TH 2796/93. *NJW* 1994, 1608 ff.

grant licenses to continue experiments on newborn non-human primates that he had applied for. The Federal Administrative Court (*Bundesverwaltungsgericht*) of Berlin regarded the authority's decision as a restriction to the basic right of freedom of research and called upon the Federal Constitutional Court (*Bundesverfassungsgericht*, the highest court in Germany) to solve this conflict between provisions of the German Animal Welfare Act and the German Constitution. The Federal Constitutional Court, surprisingly, concluded that no conflict existed. In its opinion, the Animal Welfare Act had to be interpreted in a way that it would conform to the Constitution. Such an interpretation would suggest that the authority was not entitled to make its own conclusions on the ethical justifiability of applications for animal experiments. Instead, it had to accept the applicant's reasoning, as long as this was coherent, and as long as no formal reasons stood against it.[2]

This decision had concrete consequences for the practice of licensing animal experimentation in Germany. For instance, after that decision the local competent authority of Berlin was officially advised to restrict its examination of applications for animal experiments to formal criteria.[3]

Not only to experts in the field such a restriction appeared to be in clear contradiction to basic provisions of the German Animal Welfare Act, and to Council Directive 86/609/EEC for the Protection of Animals Used for Experimental and other Scientific Purposes that Germany had to implement and enforce as a member of the European Union. In its Article 12, the Directive explicitly states "the authority shall take appropriate judicial or administrative action if it is not satisfied that the experiment is of sufficient importance for meeting the essential needs of man or animal". This provision clearly implies that the authority is expected to make its own assessment of the ethical justifiability of a proposed animal experiment.

The debate on the amendment of the Constitution in Germany

These court cases, and other court decisions that compromised animal welfare provisions, also because of the fact that they conflicted with other basic rights that were guaranteed by the German Constitution (such as the "art" of publicly killing a fish in a kitchen mixer), raised great concern not only in the German animal welfare community, but also in politicians and lawyers. They ignited a public debate that was to last for a decade. In this debate, the German Animal Welfare Federation and other animal welfare organisations lobbied and campaigned hard for insertion of a reference to animal protection in the German Constitution, so that it could be balanced against the established constitutional rights. The argument behind these activities was clear: as long as animal welfare provisions could be overruled because of the Constitution, the German Animal Welfare Act, often claimed to be one of the best legislations for animal protection in the world, would be nothing but a toothless tiger (Hobe, 1998).

On the other hand, the scientific community, in Germany and also abroad, drew up a horror scenario, prophesying that any reference to animal welfare in the Constitution would hamper scientific progress in Germany, and lead to emigration of researchers and scientific institutions, see e.g. a *Nature Neuroscience* editorial (Anon., 2002): "The likely result will be great damage to German biomedical research." These extreme reactions of researchers in Germany were not always understood well by scientists outside of Germany, as the German legal system was unique worldwide anyway, see e.g. an *ALTEX* editorial (Balls, 1999): "How can any groups in our societies expect to be allowed freedom to do what they want, unless that freedom is expressed in an acceptable ethical framework?"

The amendment of the German Constitution – wording and implications

In the year 2002 the German Constitution was amended. The amendment concerned Article 20a, which before 2002 read: "The state takes responsibility for protecting the natural foundations of life in the interests of future generations." The legal interpretation of this article had been that "life" was to mean "human life", therefore animals were not addressed in this article before 2002. The amendment of Article 20a basically consisted of the addition of the words "and animals" to the clause, which now reads:

"Mindful also of its responsibility toward future generations, the state shall protect the natural bases of life *and animals* by legislation and, in accordance with law and justice, by executive and judicial action, all within the framework of the constitutional order."

The legal interpretations of this amendment generally concluded that there were three main aspects regarding its implications (see for example Caspar und Schröter, 2003):

- Animal welfare had become a "state goal" – a matter that requires consideration when the government formulates new legislation, or existing law is interpreted by authorities and courts.
- The amendment does not grant individual rights to animals.
- Instead, it provides a legal basis for weighing animal protection measures against human interests in matters such as research and teaching.

The reality today: Court cases after 2002

To analyse whether these interpretations are reflected in legal practice, it is worthwhile to look at a court case from the year 2003, when the University of Marburg filed a lawsuit against the local competent authority's decision to not grant licenses for experiments on rats in the context of research on drug-induced pathophysiology of weight regulation.

The Administrative Court of the City of Giessen rejected the University's lawsuit and based its decision on the amended

[2] Bundesverfassungsgericht (BVerfG), Urteil vom 20.06.1994, AZ 1 BvL12/94. *Natur und Recht* 1995, Heft 3, 135-137.
[3] Verwaltungsgericht (VG) Berlin, Urteil vom 07.12.1994, AZ VG 1 A 232/92.

Constitution: After a reference to animal welfare had been inserted into the Constitution, the local authorities had the right and the duty to perform their own ethical evaluation. The court also made it clear that the authorities had the duty to reject applications if provisions of the Animal Welfare Act (they referred to indispensability, ethical justification) were not met.[4] The University of Marburg then appealed to a higher court to revoke that decision. That court, the Supreme Administrative Court (*Verwaltungsgerichtshof*) of the City of Kassel rejected the appeal in 2004. It based its decision on the fact that, in its opinion, the applicant was unable to demonstrate the indispensability of the proposed experiment.[5]

The Court also directly referred to the new legal situation after the change of the Constitution (see above).

The reality today: A survey among members of ethics committees

The Animal Welfare Academy intended to analyse the reality of regulation of animal experiments after the Constitution's amendment. That reality should be reflected in the work of the licensing authorities and ethics committees. Therefore it is presently undertaking a survey among licensing authorities and members of ethics committees on selected aspects of the workings of the ethical evaluation process. This is being done in co-operation with a Ph.D. study at the University of Tuebingen. The Animal Welfare Academy had already undertaken surveys among members of ethics committees in Germany in the years 1989 and 1995 (Rusche, 1997; Gruber and Kolar, 1997).

The new survey will be based on questionnaires that will be sent to all local competent authorities for factual information, and to members of ethics committees for their individual assessment of the situation. Already, a pilot study has been completed: telephone interviews with animal welfare members of five (out of 33 in total) different ethics committees across the country have been carried out on both aspects (tab. 1).

[4] Verwaltungsgericht (VG) Gießen, Urteil vom 13.08.2003, AZ 10 E 1409/03.
[5] Verwaltungsgerichtshof (VGH) Hessen, Kassel, Urteil vom 16.06.2004, AZ 11 ZU 3040/03.

Summary of results from telephone interviews with members of ethics committees

- The local competent authorities have not taken any substantial measures in reaction to the change in the Constitution.
- The ethics committees' work has not changed.
- At least some applications (still) lack an appropriate ethical justification.
- In a substantial number of cases, the task to evaluate whether an experiment is at all justifiable for the proposed research goal is (still) not performed.

If the results of the telephone interviews are confirmed in the overall survey, the conclusion must be that implementation of the state goal "animal welfare" is evidently unsatisfactory.

Discussion

From the analysis of court decisions before and after the year 2002, when the change in the German Constitution took place, there is an indication that the change has had a significant impact on the way issues of licensing of animal experiments are handled by courts in Germany. Their decisions now follow the amended Constitution's principle that constitutional rights must be balanced against animal welfare requirements, and that freedom of research can no longer be regarded a right that outweighs any regulations in the field of animal welfare.

However, what remains unchanged is the fact that researchers (or institutions) can still go to court to question rejections of applications by the licensing authority. On the other hand, there is no legal provision in place to allow questioning of the approval of an animal experiment by the authority. A real balance would require such a provision - for ethics committees, animal welfare organisations and others (see also Caspar und Schröter, 2003).

The number of telephone interviews with members of ethics committees carried out in the pilot study does not allow for definite postulations, however, regarding the questions selected for this publication, the answers received give clear indications.

From these interviews it became evident that the practice of licensing of animal experiments does not appear to have changed to a satisfactory extent. There are at least some local authorities that never reject an application. From previous experiences the conclusion can be drawn that this is at least partly

Tab. 1: Telephone interviews with animal welfare members of five (out of 33 in total) different ethics committees across Germany

Questions	Answers yes	no
Has your competent authority informed your ethics committees members that a new legal situation exists after the change in Article 20a of the German Constitution?	2	3
Have you observed any changes in the evaluation of applications for animal's experiments after the adoption of animal welfare into the German Constitution in August 2002?	0	5
Have you observed any changes in the work of the competent authority after inclusion of animal welfare into the German Constitution?	0	5
Do you receive applications that contain standard arguments to answer questions regarding the ethical justifiability of the proposal?	4	1
According to your opinion, do the present circumstances allow for sound ethical evaluation?	2	3

due to the fact that the composition of ethics committees does not provide for a balance between interests in using animals and interests in protecting them. In ethics committees in Germany only one third of the members represent animal welfare whereas two thirds normally come from the scientific community.

Independent of these analyses, the fact has to be mentioned that an ethical evaluation of animal experiments carried out for regulatory purposes is still not foreseen in Germany. Such experiments only need to be reported to the authorities, and are never seen by ethics committees. A true balancing of costs for animals and benefits for humans would need to include an examination of the need for substances/products (Kolar, 2000).

Another worrying problem when talking about balancing animal welfare against human benefit is the fact that there is still no defined minimum standard for the benefit. This is to say that no research goal is insignificant enough, or, in the case of toxicological testing, no product is unessential enough to set up a legal requirement against animal experiments for that purpose. However, there are some exemptions to this. For example, in the European Union, legal measures have been taken to exclude animal testing of cosmetic products and ingredients (Anon., 2003; Ruhdel, 2004). The German Animal Welfare Act also prohibits animal experiments carried out for the development or testing of weapons, ammunition, tobacco products, and detergents. There are historical reasons why these product categories have gained specific attention, but from an ethical point of view there are other product categories that would require equal consideration. Similarly categories of academic research aims have to be examined regarding their capacity to justify animal experiments.

Like no minimum standard exists for the benefit, there is no absolute (defined) limit to (the suffering in) animal experiments, i.e. the cost. No procedure is painful or unbearable enough to be excluded by legislation. There are voluntary restrictive policies, such as in Switzerland, where the Swiss Academy of Sciences has set up guidelines that include a renunciation of extremely painful animal experiments, independent of the importance of the gain of knowledge they would promise (1995). However, from an animal welfare point of view, there is a need for legal provision in this context.

There are many other issues in animal research that need to be considered and regulated appropriately, if the state-goal animal welfare is to be taken seriously by decision-makers in Germany. As a matter of priority, however, the existing regulations of the German Animal Welfare Act need to be enforced. This seems to require immediate action by the licensing authorities and ethics committees, in particular with regard to the need to give more weight to the protection of animals when there are doubts about the indispensability of the proposed research. If at this level no significant shift of the cost-benefit balance towards animal protection can be observed, other levels of governmental administration must address this issue.

References

Anon. (1986a). Council Directive 86/609/EEC of 24 November 1986 on the approximation of laws, regulations and administrative provisions of the Member States regarding the protection of animals used for experimental and other scientific purposes. *Official Journal of the European Communities L358*, 1–13.
http://europa.eu.int/comm/food/fs/aw/aw_legislation/scientific/86-609-eec_en.pdf

Anon. (1986b). *European Convention ETS 123 for the protection of vertebrate animals used for experimental and other scientific purposes*. Strasbourg: Council of Europe.
http://conventions.coe.int/Treaty/en/Treaties/Html/123.htm

Anon. (2002). Constitutional protection for animals. *Nature neuroscience editorial*, 5/7.

Anon. (2003). Directive 2003/15/EG of the European Parliament and of the Council of 27.02.2003 amending Council Directive 76/768/EEC on the approximation of the laws of the Member States relating to cosmetic products. *Official Journal of the EU L66*, 26-35.

Balls, M. (1999). Editorial. *ALTEX 16*, 1.

Bundestag (1998). Neufassung des Tierschutzgesetzes vom 25. Mai 1998. *Bundesgesetzblatt 1093, Teil I Nr. 30, G 5702*, 1105-1120.

Caspar, J. und Schröter, M. W. (2003). *Das Staatsziel Tierschutz in Art. 20a GG. Rechtsgutachten im Auftrag des Deutschen Tierschutzbundes*. Bonn.

Gruber, F. P. and Kolar, R. (1997). Animal test advisory commissions: ethics committees in Germany. In L. F. M. van Zutphen and M. Balls (eds.), *Animal Alternatives, Welfare and Ethics. Developments in Animal and Veterinary Sciences 27* (373-376). Amsterdam et al.: Elsevier Science B. V.

Hobe, S. (1998). Tierversuche zwischen Tierschutz und Forschungsfreiheit. Verwaltungs- und verfassungsrechtliche Rahmenbedingungen für Tierversuche. *Wissenschaftsrecht 31*, 309-332.

Kolar, R. (2000). Die Abwägung der ethischen Vertretbarkeit von Tierversuchen: Theorie und Praxis. *ALTEX 17*, 227-234.

Ruhdel, I. (2004). EU sales ban on new cosmetics tested on animals: Impact on alternative methods. WTO implications and animal welfare aspects. In The three R's at the beginning of the 21st century. Proceedings of the Fourth World Congress on Alternatives and Animal Use in the Life Sciences. 11-15 August 2002, New Orleans, Louisiana, USA. *ATLA 32*, *Supplement 1*, 365-371.

Rusche, B. (1997). Erste Ergebnisse über eine neue Umfrage bei den Mitgliedern in beratenden Kommissionen nach §15 Tierschutzgesetz in der Bundesrepubik Deutschland. In H. Schöffl ,H. Spielmann, H. A. Tritthart, K. Cußler, A. F. Goetschel, F. P. Gruber, Ch. A. Reinhart (Hrsg.), *Forschung ohne Tierversuche 1996* (387-394).Wien, New York: Springer-Verlag.

Swiss Academy of Medical Sciences & Swiss Academy of Sciences (1995). *Ethical Principles and Guidelines for Scientific Experiments on Animals* (2nd ed.).
http://www.springeronline.com/sgw/cda/pageitems/document/cda_downloaddocument/0,11996,0-0-45-137181-0,00.pdf

Note: All referenced websites were accessed on Sept. 2005.

Correspondence to

Roman Kolar
Animal Welfare Academy
German Animal Welfare Federation
Spechtstr. 1
85579 Neubiberg, Germany
e-mail: roman.kolar@tierschutzakademie.de

Strategies to Reduce Animal Testing in US EPA's HPV Programme

Chad B. Sandusky[1], Megha Even[1], Kristie Stoick[1] and Jessica Sandler[2]
[1]Physicians Committee for Responsible Medicine, Washington, D.C., USA
[2]People for the Ethical Treatment of Animals, Norfolk, VA, USA

Summary
The High Production Volume (HPV) programme was launched by the US Environmental Protection Agency (EPA) in 1998. To reduce animals killed, the animal welfare community negotiated basic principles set forth in a letter from EPA to HPV participants, and reiterated in the Federal Register (2000). After five years reviewing more than 370 test plans, the outcome of this effort is disappointing. However, successful strategies were developed by these reviewers and in collaboration with conscientious companies to reduce testing and still meet the Screening Information Data Set (SIDS) requirements. These strategies are explored as they might apply to future testing programmes.

Keywords: high production volume (HPV), three Rs, chemical testing, toxicity testing, animal testing

Introduction and historical overview

The EPA's HPV programme has been in progress for five years, and was envisioned to involve commitments for 2,200 chemicals by 400 companies. Submitters were to assess existing hazard data and "data gaps", and propose a plan to fill these perceived gaps. Originally designated as a "voluntary" programme, since its inception hundreds of test plans have been submitted, many of which propose animal tests to complete the SIDS base set of Tier I data requirements. The HPV programme was modeled after a similar programme administered by the OECD, i.e., the Task Force on Existing Chemicals. In the U.S. programme, the SIDS is considered a minimum for hazard evaluation and hazard is stressed over considerations of potential exposure. While each of the 3Rs (replacement, reduction, refinement) is available to HPV participants, they have frequently been ignored and/or followed to varying degrees. This current paper provides an update to a previous presentation at the Fourth World Congress on Alternatives and Animal use in the Life Sciences, in which the US HPV programme was critiqued shortly after its implementation (Nicholson et. al., 2004).

Screening information data set (SIDS)

The Tier I SIDS data requirements which use animals are provided in table 1, along with the corresponding OECD Test Guideline number and the number of animals used for each test.

The numbers of animals used per test can vary depending on exact study design, but the total number killed for a complete data set ranges from 750-870 mice, rats, and fish. Although an exact number is not possible to calculate, we have estimated that since the inception of the HPV programme, upwards of 150,000 animals have already been killed (through April 2005).

Animal welfare guidance and principles

When animal protection organisations became aware of the proposed programme, they maintained that the programme objectives, primarily the protection of human health, would not be met and that the cost in animal lives would be exorbitant. Through the White House, they also negotiated with stakeholders (EPA/Environmental Defense/American Chemistry Council) to put minimal animal welfare principles in place. The result was

Tab. 1: Screening Information Data Set (SIDS)

Test	OECD TG	# ANIMALS
Acute oral toxicity	425	3 - 10
Repeat Dose (28-day and 90-day) toxicity	407, 408	40 - 65
Combined reproductive/developmental toxicity	421	675
Combined repeated dose/reproductive/developmental toxicity	422	675
Acute toxicity to fish	203	40 - 120

the "October 1999" agreement letter, sent by EPA to programme participants (Wayland, 1999). These principles were subsequently published in the *Federal Register* (2000). The main tenets of these principles encouraged the following:

- The use of *in vitro* genetic toxicity testing rather than *in vivo* (unless impossible).
- Maximising the use of existing data.
- The use of weight-of-evidence and avoid "checklist toxicology."
- The use of Structure Activity Relationships (SAR) to form chemical categories.
- No terrestrial testing (e.g., birds, etc).
- No new dermal testing (generally).
- No sub-chronic or reproductive toxicity testing on closed system intermediates.
- Special considerations for chemicals which have been previously determined to be GRAS (Generally Recognised As Safe).
- The use of validated non-animal tests as they become available and the delay of certain testing until some non-animal methods were in place.

The goal of these animal welfare principles was to minimise animal use, while still meeting the stated hazard identification goals of the programme. Our aim was to assist companies in avoiding check-the-box toxicology to fulfill the basic SIDS data set. If the recommended generalised principles of the October 1999 letter were indeed followed, the result would be a reduction in the numbers of animals killed under the HPV programme. Each of the 3R principles (replacement, reduction and refinement) was available to HPV participants. Reductions in the numbers of animals could be accomplished, for example, by using categories of chemicals to maximise existing data or by using established OECD combined protocols such as the OECD TG 422, a combined repeat-dose, reproductive, and developmental toxicity screen, instead of three separate tests to fulfill the endpoints. Refinements to tests involving animals included the use of OECD TG 425 and cytotoxicity tests instead of the traditional LD_{50}. Finally, replacing animals completely was possible in some cases, such as the use of *in vitro* genetic toxicity tests rather than *in vivo*. The principle of "thoughtful toxicology," outlined in the October 1999 letter, provided an overarching opportunity for companies to carefully analyse existing data and decide whether additional animal tests would provide information that would be useful or relevant and to avoid such testing where it would not.

What went wrong with implementation of the HPV programme?

Once initiated, it became clear that the sponsors of HPV test plans often failed to follow even the minimal guidance offered above. The guidelines were not enforceable, and there is still no mechanism in 2005 to ensure that animal welfare guidelines are followed. In many cases, companies duplicated testing unnecessarily by conducting animal tests that had already been conducted but were conducted prior to implementation of Good Laboratory Practices (GLP), or by failing to coordinate efforts with other companies sponsoring similar chemicals. In other instances, companies did not use existing published data, individually or in conjunction with other data (in a weight-of-evidence approach), to avoid new animal testing. Often times, sponsors would fail to show relevance, such as proposing acute fish toxicity tests on water-insoluble chemicals. In many cases, when it was clear that a test was not needed for HPV, the study was proposed in "anticipation" of future data requirements, primarily Registration, Evaluation and Authorisation of Chemicals (REACH). Some companies refused to use combined protocols, sometimes doubling the number of animals killed under their test plans. Even in obvious replacement opportunities, such as the use of *in vitro* genotoxicity tests, there was an inconsistent application of the principle. In its responses to test plan proposals, EPA itself frequently failed to follow or encourage sponsors to follow basic animal welfare guidance.

HPV since 2000

Scientists at PCRM and PETA have reviewed approximately 376 test plans through August 2005, representing both individual chemicals and small to large groups of chemicals. According to our figures, a full 50% of the test plans proposed from 2000 through 2002 called for animal testing, another 50% of the test plans submitted in 2003, 45% of the test plans in 2004, and 33% of the test plans submitted through May 2005 proposed new animal testing.

These test plans account for more than 150,000 animals used to date. Importantly, these figures do not include animal tests requested by EPA above and beyond those proposed in the original test plans. It is noteworthy that after more than five years and 150,000 animals killed, no additional protections have been implemented to protect human health or the environment as a result of the HPV programme. Importantly, hazard data being generated offer little in the way of assessing human risk in that exposure characterisations are discouraged and to some degree specifically excluded from the programme. Thus, there is no context to assess the large amount of hazard information being generated by the sponsors.

Additional strategies to reduce animal use in the HPV programme

In the process of reviewing hundreds of test plans, additional strategies have been developed to supplement those envisioned in the original October 1999 letter. Some of these are extensions of the original recommendations, e.g., "common sense" toxicology, identification of duplicative testing and/or overlooked data, promoting stronger weight-of evidence approaches, etc. In addition, the wise use of resources has been stressed (a full SIDS data set may cost up to $400,000 USD) as well as encouraging companies to resist regulatory pressures when testing does not make sense. In addition to the existing guidelines, and based on an extensive review of HPV test plans, additional animal welfare principles are described below.

- **Rapid Hydrolysis of Parent Chemical.** The parent chemical need not be tested in animals if it hydrolyses to well-characterised products in an aqueous environment at low pH. A bench study at stomach pH may be used to determine rate of hydrolysis and hydrolytic products. Existing data on the hydrolysis products may then be used to meet SIDS endpoints without additional testing.
- **Acidic/Corrosive/Irritating Materials**. These are usually strong acids; they may be completely ionised in aqueous environments and are expected to cause localised, corrosive effects in the GI tract. Results from animal tests will be confounded by the corrosivity of the chemical and mammalian testing would not yield meaningful results. Animal tests using such material are particularly painful.
- **Highly Reactive Materials.** These chemicals are highly reactive to air and/or water as demonstrated by physical/chemical data. Mammalian and ecotoxicity testing with these types of chemicals is not feasible.
- **Gases.** Primary concerns with these chemicals are flammability, explosivity at test levels, and/or insolubility in water. Many are asphyxiants, some are minimally toxic and rapidly excreted, so additional testing may not be feasible or will not yield meaningful results.
- **Complex Mixtures.** The product is a mixture from different manufacturing processes and/or waste streams. Additional testing with a variable mixture may not provide useful information and existing data on major constituents may be sufficient to fill SIDS endpoints.
- **Weight-of-Evidence**. Additional testing for reproductive toxicity can be eliminated if histopathology data on reproductive organs from a 90-day subchronic study are available, in combination with a negative developmental study. This guidance is provided in the Manual for Investigation of HPV Chemicals OECD Secretariat (SIDS Manual, 2004). In some cases, traditional reproduction/developmental studies are not required if existing data from other studies, such as 2-year cancer bioassays, have evaluated reproductive and developmental parameters. A separate developmental study is not required if data exists from one- or two-generation reproduction studies.

These additional strategies have been employed successfully in the US HPV programme, and have resulted in saving thousands of animals.

Implementation and implications

In order to implement these strategies, much time is spent by PCRM and PETA reviewing each test plan, conducting internet data searches, submitting detailed comments during the public comment period, and finally, contacting individual companies to discuss opportunities to eliminate or at least reduce animal testing. We encourage companies to submit revised test plans, and we offer support to those that have already used creative and well developed strategies that reduce testing (sometimes in the form of letters to the EPA). Continued review and comment will hopefully result in future opportunities to reduce animal testing further, both in the impending REACH programme and the recently announced "Extended HPV programme," planned for a January 2006 initiation and running through 2010.

References

Federal Register, Vol. 65, No. 248, December 26, 2000, pp. 81686-98. Accessible through the World Wide Web at: http://www.epa.gov/chemrtk/frnotics.htm.

Nicholson, A., Sandler, J. and Seidle, T. (2004). An evaluation of the US High Production Volume (HPV) Chemical-testing Programme: A Study in (Ir)Relevance, Redundancy and Retrothinking. *ATLA 32, Supplement 1*, 335-341, Fourth World Congress.

SIDS Manual. Accessible through the World Wide Web at: http://webdomino1.oecd.org/comnet/env/tf-ec.nsf.

Wayland, S. (1999). Letter to HPV programme participants, 2pp. Washington, DC, USA: US Environmental Agency. Accessible through the World Wide Web at: http://www.epa.gov/chemrtk/ceoltr2.htm.

Acknowledgements

The Doris Day Animal League and the Humane Society of the United States contributed to the on-going test plan review, with either expertise and/or financial support.

Correspondence to

Chad B. Sandusky
Physicians Committee for Responsible Medicine
5100 Wisconsin Ave., N.W., Suite 400
Washington, D.C., 20016
USA
e-mail: csandusky@pcrm.org

LD_{50} Testing of Botulinum Toxin for Use as a Cosmetic

Martin L. Stephens[1] *and Michael Balls*[2]
[1]Humane Society of the United States, Washington, DC, USA; [2]FRAME, Nottingham, UK

Summary
In 2003, the Fund for the Replacement of Animals in Medical Experiments (FRAME) revealed that the potency of botulinum toxin for use as a popular wrinkle treatment is assessed by using the LD_{50} Test. The endpoint in this mouse-based testing is death through suffocation. In 2004, The Humane Society of the United States (HSUS) sought to work with Allergan, the US-based manufacturer of Botox® Cosmetic, and the U.S. Food and Drug Administration (FDA), on ways to refine, reduce, and replace this LD_{50} testing. This article summarises The HSUS's campaign in the United States and provides an update on FRAME's continuing efforts in Europe.

Keywords: Allergan, LD_{50}, Botox, botulinum toxin, Dysport, FDA, potency

Introduction

Botulinum toxin, produced by various bacteria, including *Clostridium botulinum*, is one of the most powerful biological toxins known, which blocks transmission of nerve impulses to muscles. Eight subspecies of the bacterium produce seven distinct types of toxin (types A-G), which act through different mechanisms. Food-borne botulinum toxin causes Botulism.

Several companies use botulinum toxin as the active ingredient in therapeutic products for treatment of conditions such as cervical dystonia, strabismus, blepharospasm, and hyperhidrosis. For example, Ipsen Limited UK markets a botulinum toxin Type A product, Dysport®, while the Allergan Corporation, based in the United States, markets two botulinum toxin Type A products: Botox®, for therapeutic applications, and Botox® Cosmetic, the popular wrinkle treatment – a cosmetic application.

Botulinum toxin is produced for commercial application in fermentation batches seeded with the bacteria. The standard method for assessing the potency of botulinum toxin batches is a mouse LD_{50} Test (Bottrill, 2003). In this procedure, mice are sorted into dose groups, given a single injection of toxin, and monitored over 3-4 days. Death is the endpoint, which results from suffocation through paralysis of the diaphragm musculature. Although the precise details are not available, over 100 mice are used per test, and the mouse testing is carried out up to three times prior to batch release. Calculations from the test data yield an LD_{50} value (the dose which would kill half the number of animals in a test group), which is then standardised as one "mouse unit".

The LD_{50} testing of botulinum toxin products runs counter to three trends in the application of the Three Rs of replacement, reduction, and refinement, and in animal welfare generally. First, the use of the LD_{50} test is being phased out worldwide. This was symbolised most dramatically in the field of industrial chemicals, when, in 2002, the Organization for Economic Cooperation and Development deleted the LD_{50} Test (its Test Guideline 401) from its Health Effects Test Guidelines (OECD, 2002). Second, the use of death as an endpoint is the *bête noire* of the growing field of humane endpoints (Olfert et al., 1998; OECD, 2000; ILAR, 2000). The third trend, applicable to LD_{50} testing of Botox Cosmetic, is the phasing out of animal testing of products with a cosmetic use. For example, in 2004, the European Union banned all forms of animal testing of cosmetic products (Europa, 2003).

In 2003, the Fund for the Replacement of Animals in Medical Experiments (FRAME) drew attention to the issue of LD_{50} testing of botulinum toxin products with the publication of an exposé entitled "Growing Old Disgracefully ..." (Bottrill, 2003; Balls, 2003). The Humane Society of the United States (HSUS) then took up the challenge in the USA. FRAME focuses its efforts on the European scene, and The HSUS on the US scene, but we are pleased to have this opportunity to show that we work together, and to provide brief updates on earlier assessments of alternatives to the mouse LD_{50} testing of botulinum products (Bottrill, 2003; Balls, 2003) and on the FRAME campaign, as well as a summary of the HSUS campaign.

Background

An update on alternatives for the potency assessment of botulinum toxin products

Table 1 summarises some of the potential alternatives to the mouse LD_{50} Test for assessing the potency of botulinum toxin products. Much of this information is taken from Bottrill's 2003 review. Potential refinements include mouse-based methods that assess local paralysis either *in vivo* or *ex vivo*, in contrast to systemic paralysis *in vivo*, as in the LD_{50} assay. Potential replacements target the specific molecules involved in nerve transmission that are disrupted by the various types of botulinum. For example, the SNAP-25/endopeptidase assay (Ekong et al., 1997) assesses *in vitro* the extent to which botulinum toxin Type A disrupts the activity of synaptosomal-associated protein of molecular mass 25kDa (SNAP-25), a molecule with a critical role in transmitting nerve signals.

In their recent monograph on testing botulinum toxin, the influential European Pharmacopeia (Anon., 2005) recognised the potential of these methods to substitute for the mouse LD_{50} test, stating that: "After validation with respect to the LD_{50} assay (reference method), the product may also be assayed by other methods that are preferable in terms of animal welfare", including the *in vitro* endopeptidase assay, the *ex vivo* assay using the mouse phrenic nerve diaphragm, and the "mouse bioassay using paralysis as the endpoint."

The European Pharmacopeia monograph was published prior to the publication of the Endopep-MS assay *in vitro* method (tab. 1), which has the potential advantage of permitting assessments of the potencies of all of the types of botulinum toxin.

An update on the FRAME campaign

FRAME is concerned that Ipsen Limited UK continues to use the mouse LD_{50} test to measure the potency of Dysport, and urges the Home Office (the British Government department responsible for the control of animal experimentation) to do more to bring about an end to the animal testing of botulinum toxin products for clinical and/or cosmetic use, and considers that the Government should close the loophole which permits botulinum toxin destined to be used for cosmetic purposes to be tested in animals, despite the ban on testing cosmetic products in the UK. The claim is that botulinum toxin is only officially licensed (and therefore tested) for clinical use, and its use for cosmetic purposes involves a private contract between a physician and a patient, at their own risk.

FRAME applauds the efforts of the UK national control agency, the National Institute for Biological Standards and Control (NIBSC), to develop and use refinement and replacement alternatives. The NIBSC uses *in vitro* methods on a routine basis, and only uses a non-lethal *in vivo* test when, rarely, the results of an *in vitro* test are inconclusive or close to pass/fail specifications.

FRAME is also encouraged by the effort being put by the NIBSC into the development of methods which could totally obviate the need for animal testing (eg, Ekong et al., 1997), and also that Ipsen Limited UK are working with the NIBSC and others to develop suitable batch release tests.

Meanwhile, at the European level, the European Centre for the Validation of Alternative Methods (ECVAM) and the European Directorate for the Quality of Medicines (EDQM) are working together and with others to review what progress is being made in applying the Three Rs to botulinum toxin testing and to assist in moving forward.

Summary of HSUS campaign

The HSUS campaign focuses exclusively on the testing of Botox Cosmetic by its manufacturer, Allergan, Inc., based in California. Botox Cosmetic wrinkle treatment is the most common cosmetic procedure in United States, with 2.8 million treatments carried out in 2004 (Allergan, 2005; ASAPS, 2005), accounting for 40% of net Botox sales or $295M. The HSUS regards the LD_{50} testing of products for cosmetic use as unacceptable, and seeks to hold Allergan accountable.

The strategy was to first seek to work *with* Allergan, and only if that approach failed, would The HSUS seek to pressure the company from the outside. Three things were sought from Allergan:

1. public disclosure of the details of its current potency testing of Botox Cosmetic;
2. public disclosure of the details of its current efforts to develop alternatives to the mouse LD_{50} testing of Botox Cosmetic; and
3. adoption of a well-funded and publicly available plan to rapidly end the LD_{50} testing of Botox Cosmetic.

For several months, beginning in January 2004, The HSUS engaged in cordial, but largely fruitless, communication with Allergan. The company communicated with The HSUS only through its legal staff. At The HSUS's request, the company met with Alan Goldberg, Director of the Johns Hopkins Center for Alternatives to Animal Testing (CAAT), to discuss potential CAAT/Allergan collaboration on alternatives, but the company never followed up this suggestion.

Allergan did confirm that the company uses the LD_{50} Test to assess the potency of Botox Cosmetic, and claimed to have an active alternatives program to replace this testing. However, the company provided few details either of its current testing practices or of its alternatives efforts. In its defense, Allergan noted that Botox Cosmetic and its sister product, Botox, share the same active ingredient, botulinum toxin Type A, so LD_{50} testing for the two products is inextricably linked and testing for cosmetic purposes cannot be cleanly separated from testing for therapeutic purposes. Allergan also noted the international regulations calling for the LD_{50} testing of botulinum toxin products.

The HSUS took note of these claims, but concluded that they collectively failed to justify the company's secrecy concerning its testing and alternatives practices. In The HSUS's view, any company that is making $300M a year on the backs of suffocating animals deserves to be held publicly accountable for working towards an urgent solution, especially in the context of a cosmetic application.

Tab. 1: Promising alternative methods to the mouse LD_{50} test for assessing the potency of botulinum toxin products

Name of Test	System	Endpoint	Duration	Reference
Mouse hind-limb assay	*in vivo*	local paralysis	2 days	Pearce et al., 1995
Abdominal ptosis assay	*in vivo*	local paralysis	< 1 day	Takahashi et al., 1990
Mouse phrenic nerve-hemidiaphragm system	*ex vivo*	muscle contraction	< 1 day	Bigalke et al., 2001
SNAP- 25/Endopeptidase assay	*in vitro*	molecular disruption of nerve transmission	< 1 day	Ekong et al., 1997
Endopep-MS	*in vitro*	molecular disruption of nerve transmission	< 1 day	Boyer et al., 2005

Consequently, The HSUS decided to implement the second, conditional part of its strategy towards Allergan, namely, applying public pressure to the company. Beginning in October 2004, The HSUS began issuing calls to its members and constituents to urge the company to work with The HSUS on rapidly replacing LD_{50} testing for Botox Cosmetic. The calls were issued through the organisation's electronic and hard-copy publications. In response to one appeal, thousands of e-mails to Allergan compelled the company to shut down an e-mail account.

Since Allergan refused to work with The HSUS or to disclose information about its testing and alternatives practices, The HSUS turned to the U.S. Food and Drug Administration (FDA), which had approved Botox Cosmetic and Botox. The FDA regulates these products as pharmaceuticals, and now oversees their manufacture and sale. The HSUS was specifically interested in information about the potency testing currently required or encouraged for these products. It was hoped that the agency could help answer several key questions, including the following.

1. What are the current testing practices?
2. Does the FDA require or encourage these practices?
3. How have these practices changed over the years?
4. What is the FDA itself doing to promote LD_{50} alternatives?

In 2004, The HSUS filed two Freedom of Information Act requests with the FDA, in order to obtain the sought-after information, but the agency was largely unresponsive. Consequently, the HSUS initiated legal action in 2005 to obtain the requested documents. This legal action is still pending.

Discussion and conclusions

FRAME, The HSUS, and similar organisations engage in advocacy of the Three Rs, because they want to accelerate the pace of progress in the development, validation, and implementation of methods to replace, reduce and refine animal experimentation and testing. In the case of the potency testing of botulinum toxin products, it is clear that some progress on alternative methods had been made prior to the launch of the FRAME and HSUS campaigns. However, it is clear that an unknown, but undoubtedly large number of mice were being used, and are still being used, in painful and lethal procedures for the testing of products destined for use for cosmetic, as well as for clinical, purposes. We are prepared to give Allergan, Ipsen Limited UK, and other manufacturers of botulinum toxin products the benefit of the doubt, by accepting that they are seeking to contribute to progress in the right direction.

Given this concession, some might conclude that our advocacy efforts are misplaced. We would disagree. As outlined above, the LD_{50} testing of botulinum toxin products in general, and of products for cosmetic use in particular, runs counter to three trends: the phasing out of the LD_{50} test, of the use of death as an endpoint, and of any animal testing of products with a cosmetic purpose. Consequently, the continuation of such testing is particularly out of step with the times, and is therefore particularly in need of scrutiny and action. Instead of assurances that progress is being made, what is needed is a demonstration of goodwill and verifiable action on the part of the manufacturers and the agencies responsible for the registration and use of pharmaceutical and cosmetics products and for the regulation of laboratory animal experimentation. The technical challenges to developing a non-animal alternative for botulinum toxin product testing are formidable, and are best met with collaborative efforts open to scrutiny and to constructive criticism, not with alleged programs happening behind closed doors. We note for the record that none of the published studies of alternatives to LD_{50} testing of botulinum toxin products, of which we are aware, were conducted by scientists working for the manufacturers of these products.

One of the factors that has worked against the FRAME and HSUS campaigns is the limited media interest that these efforts have generated. We suspect that this stems, in part, from limited public (and media) sympathy for mice. This is unfortunate, given that the capacity of mice to suffer is similar to that in most other animals used in laboratories. We suspect that the complexity of the relevant issues also limits the media appeal of our campaigns, including the technical nature of the non-animal alternatives and the dual use of botulinum toxin production batches for both therapeutic and cosmetic purposes. Our campaigns are also hampered by the lack of publicly available details about current testing practices and alternatives efforts.

FRAME is encouraged by the attention now being paid by the Home Office to the questions we have raised, and by the work being conducted by the NIBSC, as well as by the attention now being paid to the botulinum toxin testing issue by ECVAM and the EDQM. However, having legitimately raised an important issue of great concern in relation to both the severity of animal procedures and the need for an active commitment to finding relevant and reliable replacement alternatives, FRAME will expect progress to be made and to be kept fully informed about it.

The HSUS anticipates that its legal action against the FDA will yield critical information about the botulinum toxin testing, including the numbers of animals used per test and the number of tests conducted prior to release of a given batch of product. If Allergan continues to spurn legitimate demands for information and for co-operation, The HSUS will seek to increase the public pressure on the company, in a manner consistent with the successful campaign strategy of that late American activist, Henry Spira (Singer, 1998).

Meanwhile, both FRAME and The HSUS are willing to work collaboratively with the relevant authorities, and with institutions such as ECVAM, the EDQM, the Interagency Coordinating Committed on the Validation of Alternative Methods (ICCVAM), and the NIBSC, to accelerate the pace of progress, in the confident belief that we all share the same interest in making available products which are made as safe as possible for human use, but by using modern methods and progressively reducing reliance on the traditional application of painful test procedures to laboratory animals.

References

Allergan, Inc. (2005). Allergan Reports Fourth Quarter Operating Results; Pharmaceutical Sales Increased 16 Percent

for the Fourth Quarter; Board of Directors Declares Increase of Quarterly Dividend. http://www.shareholder.com/AGN/ReleaseDetail.cfm?ReleaseID=155017.

American Society for Aesthetic Plastic Surgery (ASAPS). (2005). Cosmetic Surgery National Data Bank 2004. http://www.surgery.org/download/2004-stats.pdf.

Anonymous (2005). Botulinum toxin type A for injection. *European Pharmacopeia Monograph 5.0*, 1117-1119.

Balls, M. (2003). Botulinum toxin testing in animals: the questions remain unanswered. *ATLA 31*, 611-615.

Bigalke, H., Wohlfarth, K., Irmer, A. and Dengler, R. (2001). Botulinum A toxin: Dysport improvement of biological availability. *Experimental Neurology 168*, 162-170.

Bottrill, K. (2003). Growing old disgracefully: the cosmetic use of botulinum toxin. *ATLA 31*, 381-391.

Boyer, A. E., Moura, H., Woolfitt, A. R. et al. (2005). From the mouse to the mass spectrometer: detection and differentiation of the endoproteinase activities of botulinum neurotoxins A-G by mass spectrometry. *Anal. Chem. 77*, 3916-3924.

Ekong, T.A.N., Feavers, I. M. and Sesardic, D. (1997). Recombinant SNAP-25 is an effective substrate for Clostridium botulinum type A toxin endopeptidase activity in vitro. *Microbiology 143*, 3337-3347.

Europa. (2003). Directive 2003/15/EC of the European Parliament and of the Council of 27 February 2003 amending Council Directive 76/768/EEC on the approximation of the laws of the Member States relating to cosmetic products (Text with EEA relevance). http://europa.eu.int/eur-lex/lex/LexUriServ/LexUriServ.do?uri=CELEX:32003L0015:EN:HTML.

Institute for Laboratory Animal Research (ILAR). (2000). Humane Endpoints for Animals Used in Biomedical Research and Testing. *ILAR Journal 41*, No. 2.

Olfert, E., Bhasin, J., Latt, R. et al. (1998). *Guidelines on: choosing an appropriate endpoint in experiments using animals for research, teaching and testing*. Ottawa: Canadian Council on Animal Care (CCAC).

Organisation for Economic Cooperation and Development (OECD). (2000). Guidance Document on the Recognition, Assessment, and Use of Clinical Signs as Humane Endpoints for Experimental Animals Used in Safety Evaluation. No. 19. Paris: OECD.

Organisation for Economic Cooperation and Development (OECD). (2002). OECD Test Guideline 401 will be deleted: A major step in animal welfare. http://www.oecd.org/document/52/0,2340,en_2649_34377_2752116_1_1_1_1,00.html.

Pearce, L. B., First, E. R. and Borodic, G. E. (1995). Botulinum toxin: death versus localized denervation. *Journal of the Royal Society of Medicine 88*, 239-240.

Singer, P. (1998). *Ethics into Action: Henry Spira and the Animal Rights Movement*. Lanham, Boulder, New York, Oxford: Rowman & Littlefield.

Takahashi, M., Kameyama, S. and Sakaguchi, G. (1990). Assay in mice for low levels of Clostridium botulinum toxin. *International Journal of Food Microbiology 11*, 271-277.

Acknowledgements

The authors wish to thank Loree Talley and Robert Combes, for their assistance with this manuscript and in the associated campaigns.

Correspondence to

Dr Martin L. Stephens
Humane Society of the United States
2100 L Street, NW,
Washington, DC 20037
USA
e-mail: mstephens@hsus.org

Professor Michael Balls
FRAME
Russell & Burch House
96-98 North Sherwood Street,
Nottingham NG1 4EE
UK
e-mail: michael.balls@btopenworld.com

Session 3.5
Ethical review – good practice and outputs

The Role and Evolution of Independent Government Advisory Committees: The New Zealand Experience from 1985 to 2005

A. C. David Bayvel
Ministry of Agriculture and Forestry, Wellington, New Zealand

Summary
In the New Zealand national animal welfare infrastructure, the National Animal Ethics Advisory Committee (NAEAC), the National Animal Welfare Advisory Committee (NAWAC) and the Australian and New Zealand Council for the Care of Animals in Research and Teaching (ANZCCART) all play important and discrete roles in ensuring that the New Zealand government receives independent, broadly-based advice regarding the use of animals in science, in agriculture and for other purposes. This paper covers the genesis, legal status and management disciplines associated with the New Zealand NAEAC/NAWAC system, including the role of the scientific community, Government and NGOs.

Keywords: government advisory committees, animal welfare, animal ethics, New Zealand

Introduction

Over the last two decades, animal welfare has received increasing attention in New Zealand as a complex, multi-faceted public policy issue with important scientific, ethical, economic, cultural and social dimensions. Over this same period, New Zealand has developed a cohesive national animal welfare infrastructure, animal welfare science and ethics capability and has introduced contemporary animal welfare legislation (Bayvel, 2000; Mellor, 1999; Ministry of Agriculture and Forestry, 1999; Ministry of Agriculture and Forestry, 2000; Mellor and Bayvel, 2004). The latter takes note of both the concepts of the Three Rs of Russell and Burch (1959) and the five freedoms originally developed by the UK Farm Animal Welfare Council (Farm Animal Welfare Council, 1993).

Two independent ministerial advisory committees play a vital role in the national infrastructure, with their functions, responsibilities and membership prescribed by the Animal Welfare Act 1999. The National Animal Ethics Advisory Committee (NAEAC) deals with the use of live animals in research, testing and teaching, while the National Animal Welfare Advisory Committee (NAWAC) covers all other uses of animals, including in agriculture and in entertainment.

In the development of animal welfare policy and the setting of regulatory standards, the consultative requirements of the Animal Welfare Act provide extensive opportunity for input from both the public at large and affected and interested stakeholders. This input is complemented by high level independent advice to Government from two committees, whose membership reflects the broad spectrum of opinion which exists within any society on animal use in science, agriculture and for other purposes.

Discussion

History and background

Reid (1989) provided a very valuable account of the historical developments leading up to the establishment in New Zealand of Animal Ethics Committees (AECs) and the National Animal

Ethics Advisory Committee (NAEAC). In the 1970s and 1980s, the New Zealand scientific community had recognised the need for legislation covering the use of live animals in research, testing and teaching and, through the Royal Society of New Zealand, had worked closely with Government in formulating appropriate policy and legislation. AECs were, to an extent, modelled on existing human ethics committees and also took note of the experience gained in countries such as Sweden and Canada. In spite of some opposition, it was decided that NAEAC should be located within, and serviced by, the Ministry of Agriculture as a core Government department.

NAEAC was established in 1984 and the complementary National Animal Welfare Advisory Committee (NAWAC) held its first meeting in 1989. A third significant development, in the early 1990s, was the establishment of the Australian and New Zealand Council for the Care of Animals in Research and Teaching (ANZCCART). Unlike the two ministerial advisory committees, ANZCCART has no statutory status and is constituted as a committee of the Royal Society of New Zealand. It does, however, play an important role in encouraging informed debate on the ethical use of animals in science, promoting humane science and sponsoring valuable conferences in conjunction with ANZCCART Australia.

Roles and Responsibilities

Part 4 of the New Zealand Animal Welfare Act 1999 is devoted to advisory committees and covers in detail the purpose, functions, membership, terms of office and annual reporting requirements for both NAEAC and NAWAC.

The functions of NAEAC are:

(a) To advise the Minister on ethical issues and animal welfare issues arising from research, testing, and teaching

(b) To make recommendations to the Minister under section 3(3) (which relates to manipulation)

(c) To make recommendations to the Director-General under section 85 (which relates to restrictions on use of non-human hominids)

(d) To provide advice and information on the development and review of codes of ethical conduct

(e) To make recommendations to the Director-General concerning the approval, amendment, suspension, or revocation of any code of ethical conduct

(f) To make recommendations to the Minister concerning the setting of standards and policies for codes of ethical conduct

(g) To provide information and advice to Animal Ethics Committees

(h) To recommend, for approval by the Director-General under section 109, such persons as are, in the opinion of the committee, suitable for appointment as accredited reviewers

(i) To consider the reports of independent reviews of code holders and Animal Ethics Committees:

(j) To make recommendations to the Minister under section 118(3) (which relates to the power of the Minister to approve research or testing).

The membership of NAEAC reflects the need for the committee to possess knowledge and experience in the following areas:

(i) Veterinary science

(ii) Medical science

(iii) Biological science

(iv) The commercial use of animals in research and testing

(v) Ethical standards and conduct in respect of animals

(vi) Education issues, including the use of animals in schools

(vii) The manipulation of animals in research, testing, and teaching

(viii) Environmental and conservation management

(ix) Animal welfare advocacy

(x) Any other area the Minister considers relevant; and

The membership recognises the need for a balance between those members who are currently involved in research, testing, and teaching and those members who are not so involved.

(a) To advise the Minister on any matter relating to the welfare of animals in New Zealand, including (without limitation)

(i) Areas where research into the welfare of animals is required; and

(ii) Legislative proposals concerning the welfare of animals

(b) To make recommendations to the Minister

(i) Under section 3(3) (which relates to manipulation); and

(ii) Under sections 6 and 16 (which relate to surgical procedures)

(c) To discharge its functions under section 32 in relation to the making of Orders in Council declaring traps or devices to be prohibited or restricted traps or devices

(d) To discharge its functions under section 32 in relation to the conditions that should be attached to the sale or use of any restricted trap or restricted device

(e) To make recommendations to the Minister concerning the issue, amendment, suspension, revocation, and review of codes of welfare

(f)To promote, and to assist other persons to promote, the development of guidelines in relation to

(i) The use of traps or devices or both:

(ii) The hunting or killing of animals in a wild state

The membership of NAWAC reflects the need for the committee to possess knowledge and experience in the following areas:

(i) Veterinary science

(ii) Agricultural science

(iii) Animal science

(iv) The commercial use of animals

(v) The care, breeding, and management of companion animals

(vi) Ethical standards and conduct in respect of animals

(vii) Animal welfare advocacy

(viii) The public interest in respect of animals

(ix) Environmental and conservation management

(x) Any other area the Minister considers relevant

Schedules to the Act cover a range of administrative matters and also describe in detail the provisions applying to accreditation of independent reviewers of AECs and associated code holders. In all cases, members are appointed to committees in a personal capacity recognising their individual experience and expertise and do not play a representational or advocacy role on behalf of nominating organisations.

Policies and procedures

Initially, appointments to both committees were not time-bounded. Fixed terms of three years were, however, introduced during the 1990s for both committee members and chairs. In most cases, two terms is the maximum period served by committee members and chairs but, in exceptional cases, this can and has been extended. This policy ensures a balance of committee experience and knowledge with the introduction of "new blood" but requires careful planning in relation to succession management and committee member and chair induction and hand-over. In the case of NAEAC, the independent committee chair has historically had no involvement with the use of animals in science, while in the case of NAWAC a detailed familiarity with animal agriculture and animal welfare science has been seen as an important attribute.

Formal strategic planning was introduced for both committees in the early 1990s. The strategic plan is reviewed in detail every three years and updated annually. The agreed strategic plan gives direction to the annual operational plan which, in turn, defines specific responsibilities for individual committee members and associated resource requirements.

Annual performance reviews are conducted for each committee. These involve completion of a standard questionnaire regarding committee performance and have proved to be a useful and valued tool in identifying issues and improving committee effectiveness. Policy manuals have also been prepared for both committees to record committee policies and decisions, to assist with the induction of new members and chairs and ensure consistency of decision making.

Communications

In addition to such committees generating quality advice and recommendations and effectively managing and discharging their statutory responsibilities, it is vital that they also communicate effectively with Government and all interested stakeholders. Communication with Government is assured by annual meetings of chairs with the relevant Minister, occasional attendance by the Minister at committee meetings and the preparation of an annual report which is a legislative requirement. This annual report is a major communication vehicle to all stakeholders and is complemented by annual or bi-annual newsletters from each committee.

Other strategic communication initiatives are included in an annually updated communications plan, which includes specific activities such as workshops, conferences, press conferences and press releases.

International and domestic relationships

It is in New Zealand's strategic interest to keep abreast of international trends in the area of animal welfare and ethics and to liaise closely with organisations, institutions and other Government agencies to foster exchange of information and to undertake collaborative initiatives. In the area of animal use in research, testing and teaching, New Zealand has thus developed important relationships with ANZCCART in Australia, the United Kingdom Home Office, the International Council for Laboratory Animal Science (ICLAS), the Canadian Council for Animal Care (CCAC), Three Rs centres in Europe and North America, the European Commission and the UK Research Defence Society (RDS).

In the case of animal use in agriculture, similar important relationships exist with the UK Farm Animal Welfare Council (FAWC), the UK Department for the Environment, Food and Rural Affairs (DEFRA), the European Commission, university research groups and non-governmental organisations in Europe and North America and the USDA Animal Welfare Information Center.

Formal Government to Government interaction also takes place between New Zealand and Australia via the Trans-Tasman Animal Welfare Working Group (TTAWWG), and between New Zealand, Australia, Canada and the USA via the Quadrilateral Animal Welfare Working Group (QAWWG). The World Organisation for Animal Health (OIE) involvement in animal welfare is an important recent development on the world stage and New Zealand has participated actively in this strategically important international initiative. Such international communication has been dramatically facilitated by electronic communication and the availability of information on web sites.

On the domestic front, both NAEAC and NAWAC have established close links with the New Zealand Bioethics Council (Toi te Taiao). The purpose of the Bioethics Council is to:

-Enhance New Zealand's understanding of the cultural, ethical and spiritual aspects of biotechnology

-Ensure that the use of biotechnology has regard for the values of New Zealanders.

The New Zealand Government established the Bioethics Council in December 2002 following a recommendation by the Royal Commission on Genetic Modification. The Royal Commission wished to address public concern that decision-making was not adequately considering the cultural, ethical and spiritual dimensions of biotechnology. Submissions to the Royal Commission had highlighted the need for high-level advice to Government with community input on overarching cultural, ethical and spiritual concerns that had wider implications for society. The Bioethics Council has been set up as a ministerial advisory council that sets its own work programme and priorities. It reports to the Government through the Minister for the Environment.

To promote dialogue between politicians and key stakeholders involved in the animal welfare and ethics debate, an All Party Animal Welfare Group (APAWG) was established in 1994. This group was modelled on similar groups existing in the UK and European Union and is sponsored by the Royal Society of New Zealand (via ANZCCART), the New Zealand Veterinary Association (NZVA), the Royal New Zealand Society for the Prevention of Cruelty to Animals (RNZSPCA) and Federated Farmers of New Zealand (FFNZ).

Operational research

Research funding is made available, on a contestable basis, by the New Zealand Ministry of Agriculture and Forestry for

research studies to support policy formulation and standard setting in the area of animal welfare. As indicated in the NAWAC Annual Report (National Animal Welfare Advisory Committee, 2005), over the period 1993 to 2004, 88 studies have been completed with the majority published in peer reviewed journals and the internationally recognised MAF publication "Surveillance". Such research findings have proved extremely valuable to support both committees in their independent advisory roles and to ensure that recommendations are supported, or at least informed, by relevant science.

Conclusion

Both NAEAC and NAWAC have evolved over the last 20 years in terms of core functions and supporting management and administrative disciplines. The relationship between the two committees is formalised by the NAEAC chair also being a NAWAC member. This ensures that areas of common interest and policy parallels are addressed efficiently and effectively. Ethical debate, which has been a key element of NAEAC discussions, is now also receiving increasing attention in relation to the use of animals in (particularly intensive) agriculture. The minimisation of pain and distress will continue to be a key priority for both committees.

In the case of NAEAC, genetic modification research presents a particular challenge. The Three Rs of Russell and Burch (1959) have played a critical role over the past fifty years but the opportunity to expand the Three Rs, as described by Schuppli (Schuppli et al., 2004), will assist in addressing the challenges of the future. Schuppli proposes that "Some of the concerns that fall outside the Three Rs framework might be captured by adding a fourth "R" for responsibility. Responsibility, as proposed by the International Foundation for Ethical Research, could be elaborated to include responsibility to not violate community standards, responsibility to enhance the quality of life of animals in human care, and responsibility to reflect real costs in analyses of costs and benefits." (Schuppli et al., 2004).

In a recent paper Balls (2004) emphasises the importance of the "middle ground" i.e. "the only way forward to co-operation between people of goodwill and common sense – in the middle ground – is to secure the development and use of reliable and scientifically sound alternatives. This means rejecting the extremism on both sides." Goldberg (2004) also refers to the "troubled middle" as reflecting a societal "understanding of the need for animal-based research but a desire to ensure that pain and distress are eliminated or, at least, minimised" and the "silent middle" i.e. scientists who practise humane science, who would prefer not to use animals but who do not publicly articulate their position or commitment. (Goldberg, 2004).

Goldberg further emphasises that extreme positions reject both societal expectations and scientific need and that "the issues for the general public are transparency, accountability and humaneness". In the New Zealand context, NAEAC, NAWAC and ANZCCART are essential elements of a cohesive national infrastructure and make important and unique contributions to addressing these issues.

References

Balls, M. (2004). Editorial – Animal experimentation and alternatives in the age of extremism. *ATLA 32*, 151-152.

Bayvel, A. C. D. (2000). The New Zealand regulatory approach to the use of animals in research, testing or teaching. In M. Balls, A-M. van Zeller and M. E. Halder (eds.), *Progress in the Reduction, refinement and Replacement of Animal Experimentation* (1023-1031). Amsterdam: Elsevier Science B.V.

Farm Animal Welfare Council (1993). *Second Report on Priorities for Research and Development in Farm Animal Welfare*. London: Farm Animal Welfare Council.

Goldberg, A. M. (2004). Animals and alternatives: societal expectations and scientific need. *ATLA 32*, 545-551.

Mellor, D. J. (1999). The new Animal Welfare Act 1999 and conducting animal-based science in New Zealand. *ANZCCART News 12*, 6.

Mellor, D. J. and Bayvel, A. C. D. (2004). *Application of legislation, scientific guidelines and codified standards to advancing animal welfare*. OIE Conference Proceedings 23-25 February 2004 Paris, 249-259.

Ministry of Agriculture and Forestry (1999). *Guide to the Animal Welfare Act 1999*. Wellington: Ministry of Agriculture and Forestry.

Ministry of Agriculture and Forestry (2000). *Guide to Part 6 of the Animal Welfare Act 1999*. Wellington: Ministry of Agriculture and Forestry.

National Animal Welfare Advisory Committee (2005). 2004 *Annual Report*. Wellington: Ministry of Agriculture and Forestry.

Reid, C. S. W. (1989). *Ethics, Animal Science and the Royal Society in the Use and Welfare of Experimental Animals*. Wellington: Royal Society of New Zealand Symposium Proceedings.

Russell, W. M. S and Burch, R. L. (1959). *The Principles of Humane Experimental Technique, reprinted edition*. London: UFAW,1992.

Schuppli, C. A., Fraser, D. and McDonald, M. (2004). Expanding the Three Rs to meet new challenges in humane animal experimentation. *ATLA 32*, 525-532.

Acknowledgements

The valuable contributions made by D. J. Mellor (Massey University), G Sutherland (ANZCCART NZ), L. Carsons (MAF) and M. Short (MAF) in the preparation of this paper are gratefully acknowledged.

Correspondence to

Dr. A. C. David Bayvel
Animal Welfare Group
Biosecurity New Zealand
Ministry of Agriculture and Forestry
P O Box 2526
Wellington, New Zealand
e-mail: David.Bayvel@maf.govt.nz

Processes and Policies for Ethical Evaluation in Nordic Countries and Europe

Timo Nevalainen[1,2], Anna Meller[3] and Niina Kemppinen[3]
[1]National Laboratory Animal Center, University of Kuopio and [2]Department of Basic Veterinary Sciences,
[3]Viikki Laboratory Animal Center, University of Helsinki, Finland

Summary
In ethical evaluation of animal studies both costs and benefits are broken down to smaller elements in order to attach an ethical value judgment to each, and look for improvements in areas of concern. Ethical evaluation must focus on the refinement and reduction alternatives. Evaluation of the 6th Framework Programme applications is a truly European process: Benefits are assessed first by scientific evaluators, and applications with high scores go to ethical panel. A similar assessment of benefits should be done in local ethics committees, but they may lack the needed expertise. The Nordic Forum suggested a Cost-Benefit-Means approach for evaluation.

Keywords: ethical evaluation, policy, cost benefit assessment

Background

Special attention has been paid in European regulations and policies to the 3Rs alternatives. The European Commission Directive on the protection of animals used for experimental and other scientific purposes (86/609/EEC) (1986) states that the EU Member States must actively encourage and support the development, validation and acceptance of methods which could Replace, Reduce and Refine the use of laboratory animals (3Rs). The same is true for the Treaty of Amsterdam and the policy paper of the European Science Foundation (ESF). The same approach and policy are presented in the Commission's White Paper to the New Chemicals Directive (REACH) (2001).

The Report on Directive 86/609 (2001) by the Committee on the Environment, Public Health and Consumer Policy (2001/2259(INI)) states more specifically that an ethical and animal-welfare assessment must be carried setting limits to the level of suffering and distress to which the animals may be subjected.

Revised Directive (86/609/EEC) shall require detailed and harmonised ethical evaluation of animal studies, and will be based on cost-benefit analysis. In this analysis the likely benefits of the study are weighed against the cost – i.e. harms like pain, suffering and distress – to the animal. It can be foreseen that both commodities to be weighed have to be broken down to smaller elements in order to weigh or attach an ethical value judgment to each and then these elements can be used in the overall assessment of an animal study.

What is perhaps even more important is improving all relevant areas of concern, but particularly so that both animal welfare and good science are promoted. Whenever replacement alternatives cannot be used, ethical evaluation can and must focus on the two other alternatives, refinement and reduction and these are also fundamental elements of any harm in a cost/harm-benefit analysis. Processes and policies of ethical evaluation vary considerably in Europe, and need in many cases to be modified.

The EU Framework Programmes

The Ethical Rules for the 6th Framework Programmes (FP) proposals reiterate the application of the 3Rs principles and entail a description of the procedures adopted to ensure that the amount of suffering imposed on the animals is minimised and their welfare is guaranteed as far as possible (e.g. through improvements in experimental technique, application of humane endpoints, environmental enrichment, etc.).

According to the Ethical Rules for the 6th FP proposals, applicants should provide – at the end of the application – a summary of the main adverse effects for the animals, including those due to methods of husbandry and supply of the animals as well as the harmful effects of the scientific procedures themselves.

Evaluation of the 6th Framework Programme applications is an example of truly European process: Benefit assessment is carried out first by scientific evaluators, and applications with high scores go to ethical review panel. The ethical review panel is composed of independent experts of law, sociology, psychology, philosophy and ethics, medicine, molecular biology and veterinary science. It has a parity of scientific and non scientific members. Once the application is discussed in the ethical panel, the panel can ask for more information, require or recommend changes and modifications before negotiations between the consortium and the Commission can commence.

Limited experience from ethical panels indicates that many applicant consortia do not really know what the 3Rs are. Perhaps this is due to the fact that by the end of the application they are tired by other questions to be addressed, regard other ethical issues more important or do not care to find out. It is simply not enough to state that the consortium is committed to the three Rs principles while the question is how to implement them in their own study.

A similar assessment of benefits should be done in all cost-benefit analyses, but the local ethics committees may often lack the

needed expertise for detailed scientific scrutiny. Perhaps local ethics committees should ask for evaluations received with applications for funding from major national or European sources. This approach would leave benefit assessment of only small or new groups to the local ethics committees.

The Nordic Forum

The Nordic Forum for Ethical Evaluation of Animal Experiments, held in Helsinki 2003, was arranged to see whether a suitable scoring system for a cost-benefit analysis are available, and if any evaluation scheme could be agreed among the Nordic participants (Voipio et al., 2004). The workshop introduced the terms "Cost modifiers" or "means" which can be used to decrease the costs for the animals, or to improve the benefits from the study. The use of scoring systems for ethical assessment did not receive support since they were considered to give a false impression of objectivity. A classification of costs versus benefits into three degrees (low, medium and high) was considered the most suitable approach. Example protocols were evaluated by the workshop participants; the result revealed a large variation in scoring the degree of costs, the importance of the benefits, and the possibilities of modifying the means. Clearly, further and continuing interaction between all of the interest parties is necessary for the creation of more precise tools for ethical assessment of animal studies.

The two Rs initiative

Until now the replacement alternative has received far more attention and EU funding opportunities than the other Two Rs – reduction and refinement. An initiative – called the Two Rs Initiative – to be included into the 7th Framework Programme (2007-2013) has been submitted by the COST Action B24 "Laboratory Animal Science and Welfare" to the Commission and was endorsed by 50 European institutions and scientific associations.

Since replacement is not always possible, the animals still used should not be neglected. To help the animals and science more research on the Two Rs methods should be encouraged leading to better quality animals and to a reduction of the numbers used. The COST Action believes that even better science evolves from the application of the Two Rs.

The application of the Two Rs will also be instrumental when practising cost-benefit analysis, as will be the quality of use, care and housing of laboratory animals. They should be regarded as the means to either increase the benefits or decrease the costs in a study. This is particularly true with GM-animals and some new research methods.

The welfare and number of animals used can be regarded as essential elements in assessment of costs/harms, incurred to the animals in the study. This approach can only be successful if tackled by multidisciplinary teams, i.e. by both the study groups themselves and laboratory animal scientists. Involvement of both parties is necessary for attainment of critical mass, a prerequisite to large scale outcome. And the Two Rs means will be creditable only if they have proven 'efficacy' to animal welfare and proven 'safety' to the study.

European consortia and European funding are needed to coordinate national research activities and studies on the Two Rs, to attain European added value through implementation of the Two Rs methods and to show that the European research community practices good ethics in their daily work. By high quality research, improving animal welfare and avoiding unnecessary duplication of animal studies, a valuable contribution to sustainable economic growth is attained within the EU.

References

Commission of the European Communities (2001). *White Paper. Strategy for Future Chemicals Policy*, Brussels.

Committee on the Environment, Public Health and Consumer Policy. (2001). *Report on Directive 86/609*.

Council Directive 86/609/EEC of 24 November 1986 on the approximation of laws, regulations and administrative provisions of the Member States regarding the protection of Animals used for Experimental and Scientific purposes.

European Commission. *Ethical Rules for FP6*.

European Science Foundation. (2001). *Use of Animals in Research.*

Voipio, H-M., Kaliste, E., Hirsjärvi, P. et al. (2004). Nordic-European Workshop on Ethical Evaluation of Animal Experiments. Workshop Report on the Cost-Benefit Principle in Hanasaari, Helsinki, Finland 7-9 November 2003. *Scand. J. Lab. Anim. Sci. 31*, 251-267.

Correspondence to

Prof. Timo Nevalainen
NLAC / University of Kuopio
POBox 1627
70211 Kuopio, Finland
e-mail: Timo.Nevalainen@uku.fi

Revision of a Scale for Assessing the Severity of Live Animal Manipulations

Virginia M. Williams[1], David J. Mellor[2] and John Marbrook[3]
[1]School of Natural Sciences, Unitec, Auckland, New Zealand; [2]Animal Welfare Science and Bioethics Centre, Massey University, Palmerston North, New Zealand; [3]Auckland, New Zealand

Summary
In 1997, a severity scale to assess and record the level of welfare compromise to animals used in research, testing and teaching was introduced in New Zealand. Under this scale, the severity of procedures was expressed in terms of different categories of suffering based on numerous examples at the five levels outlined in a paper by Mellor and Reid (1994). This paper reports on a review into the operation and effectiveness of that scale and the extent to which it fulfils the purposes for which it was devised. Key features of the scale are described, including its strengths and limitations, and comparisons with other scales operating internationally are made. Recommendations regarding modification of the scale based on this evaluation are outlined, and key steps in its implementation are described.

Keywords: animal welfare, research, testing and teaching, impact scale, severity scale, adverse effects in animal research, animal suffering

Introduction

The New Zealand Animal Welfare Act (1999) requires that any person or institution wishing to manipulate live animals for the purposes of research, testing or teaching must operate under an approved Code of Ethical Conduct (CEC). The code requires that an Animal Ethics Committee (AEC) be set up so that all projects can be scrutinised, and only approved projects may proceed. Although the Animal Welfare Act itself is relatively recent, this system has in fact been in operation since 1987 as a result of the process used to implement the Animals Protection (Codes of Ethical Conduct) Regulations 1987.

One of the functions of the AEC is to carry out a cost-benefit analysis of projects, with more “invasive” or “severe” manipulations requiring greater justification for the work in terms of the anticipated benefits that accrue from it. Applicants presenting protocols to AECs must therefore predict the level of animal welfare compromise expected as a result of the proposed manipulations, although that assessment may be corrected once the manipulation has taken place if the actual degree of compromise is found to have varied from that expected.

As detailed by Bayvel (2004), the data from such assessments are collected for each approved study, accumulated on an institutional basis and, after submission to the Ministry of Agriculture and Forestry (MAF), collated as annual national figures. These provide published details of the number of animals of different types exposed in New Zealand to manipulations of different severity levels on the specified scale. Thus, a more detailed picture of annual animal use in New Zealand is provided than would be from a grand total for animal use.

In assessing welfare compromise, New Zealand uses a five-point scale devised by Mellor and Reid (1994), which was introduced in 1997. Concerns have been raised, however, about a probable lack of consistency across institutions in the way the scale has been applied. This could have resulted in both under- and over-assessment of the impact of manipulations. A tendency for the collected statistics to be misrepresented has been dealt with in part by a new requirement in the collection of statistics by MAF, which differentiates animal use according to whether it is for research, testing or teaching. Thus the relatively large numbers of animals that are used in testing required by legislation, some of which have a high degree of welfare compromise, are separated out within the total number of animals used. This confirms, for the concerned public, that experimental research or teaching rarely involves animals at high levels of "suffering”, contrary to the charge that is often levelled at researchers by animal rights protestors.

It has also become apparent that the purposes of the scale are not well understood by some of those using it, with the result that it has, in a few cases, been applied by simply using the minimal definitions (no suffering, little suffering, moderate suffering, severe suffering and very severe suffering) found on the MAF statistics collection form. Others have merely used the examples given in the full explanation provided by MAF (Anon (b), 2001) in a way that prevents the greater flexibility and judgement allowed under the five domains as intended in the original paper (Mellor and Reid, 1994).

Concerns have also been raised about whether the use of animals in biotechnology is adequately covered within the current system, with the suggestion that separate categories might be more appropriate.

Accordingly, as part of a process of continuing improvement followed since the system was first introduced in 1987, MAF commissioned research to review the rationale, terminology and

explanation of the scale and to compare it with others now used overseas, with the aim of ensuring clarity and minimising ambiguity in definitions and descriptions of purposes and differences between categories.

Purpose of grading the impact of manipulations

A severity scale for animal-based scientific manipulations is of significance to four interested parties: animal-based scientists, animal ethics committees, regulators and the public.

- *Animal-based scientists* need to assess the invasiveness of their manipulations as an integral part of seeking approval to undertake them. Evaluating invasiveness gives an indication of the harm that may be done to the animals. Conducting a harm-benefit analysis is a pivotal part of achieving approval for a proposed manipulation. This ensures that the potential adverse effects on the animals used are greatly outweighed by the benefits accrued from the work. This way of justifying the work is in compliance with the utilitarian ethical basis for using animals in research, testing and teaching.
- *Animal ethics committees* are required by the Animal Welfare Act and their institution's Code of Ethical Conduct to undertake a harm-benefit analysis for each application to manipulate animals in research, teaching and testing. This too relates to the utilitarian ethical mode of assessing the justification for such animal use as proposed in applications to each committee. It also allows members of the AEC to help to ensure that the principle of refinement, or minimisation of harm, is fulfilled.
- *Regulators* require it because it confirms that such assessments are indeed being undertaken in compliance with the Animal Welfare Act, and this can be demonstrated by the public release of annual statistics showing the range of invasiveness of approved manipulations.
- *Interested members of the public* desire knowledge of this kind in order to be reassured that (1) excessively invasive manipulations are not being conducted in a high proportion of the animals, (2) very invasive manipulations are properly justified, and (3) within the full range of severity, the majority of manipulations have been at the bottom end of the range where the impacts are very low (benign) with few negative consequences.

Key features of the current scale

- *Five levels of severity are defined.* They are O, A, B, C and X, and range from no, or virtually no, impact (Grade O) to the most severe impact that one could imagine would ever be approved (Grade X). They are currently expressed in the Animal Welfare (Records and Statistics) Regulations 1999 in terms of severity of suffering, although this was not the mode of expression employed by Mellor and Reid (1994) in their original description of the scale. They referred to different levels of "ethical cost", not suffering.
- *The predicted severity of welfare impacts is assessed comprehensively*. This is achieved by reference to *five domains of potential welfare compromise*, which relate to the nutritional, environmental, health maintenance, behavioural and mental needs of animals. The final assessment of the severity of welfare compromise refers mainly to the *mental domain*, because the overall welfare status of an animal is directly reflected in its affective or mental state. Nevertheless, consideration of the other four domains is imperative, as it helps to ensure that the welfare status of the animals to be manipulated has been assessed thoroughly in all of its dimensions.
- The predicted severity of impact determines the required level of justification. The different levels of justification were characterised by Mellor and Reid (1994). The greater the impact, i.e. the more severe the manipulation, the greater is the justification that is required for animal ethics committee approval to be given. Those justifications range from no requirement to demonstrate immediate or even long-term benefits for people or animals for Grade O manipulations, to the most exceptionally strong justifications in terms of benefits to people and/or animals for Grade X manipulations. Grade X should contribute to resolving a pressing need of great significance.

International comparisons

A paper that considered the possibility of international harmonisation of "pain and distress" classification systems (Purves, 2000) highlighted the considerable international variation in the way such systems are used.

In the present critical review, the New Zealand system was compared with nine other current scales (tab. 1), all of which were included in the Purves paper (2000), except one developed by New South Wales Agriculture in Australia (Anon (c), 2003). The Swedish system, which was legally abandoned in 1989 and only recently reinstated, was omitted.

Seven of the ten systems used classifications based only on the effect on the animal, with descriptors including discomfort, stress, pain, sorrow, harm, fear, suffering, severity degree and constraint. In addition, some of these included consideration of the duration of manipulation.

Of the other three, the Australian scale alone classifies types of procedures rather than effects on the animal, while the other two combine features of both systems: the New South Wales system outlines broad categories of intervention, differentiating between minor and major surgical and physiological challenges, while the US system, having separated out procedures causing little or no pain and distress, simply divides the remainder according to whether or not pain relief is given.

The Impact Scale

The demands for transparency made by interested members of the public will not be met if, for whatever reason, the categorisation within the scale is not accurate or meaningful, the scale does not give a balanced view of overall animal use or the scale does not support the production of meaningful statistics. The present critical review confirmed the sound basis for categorisation set in the original paper by Mellor and Reid (1994), thus

Tab. 1: Comparison of available "impact" scales

Comparison of international impact scales						
Country	Year of adoption	No. categories	No. manipulated* categories	Severity-based Classification	Other Classification	Duration acknowledged
New Zealand[1]	1997	5	5	Yes		Yes
Australia[2]		13	12		Yes (manipulation type)	
Canada[3]	1987	5	4	Yes		
Finland[4]	1986	3	3	Yes		
Germany[5]		4	4	Yes		Yes
Netherlands[6]	1979	3	3	Yes		Yes
NSW[7]	2003	9	9	Mainly	Phys v surgery, GM	
Switzerland[8]	1994	4	4	Yes		
UK[9]	1986	3	3	Yes		Mild band only
USA[10]		3	3		Yes (+/-pain relief)	

*A manipulation is:
(a) subjecting an animal to a procedure which is unusual or abnormal when compared with that to which animals of that type would be subjected under normal management or practice and which involves
(i) Exposing the animal to any parasite, microorganism, drug, chemical, biological product, radiation, electrical stimulation, or environmental condition; or
(ii) Enforced activity, restraint, nutrition, or surgical intervention; or
(b) Depriving the animal of usual care (New Zealand Animal Welfare Act, 1999).

[1] New Zealand (Mellor and Reid, 1994)
[2] Australian (Purves, 2000)
[3] CCAC – Categories of Invasiveness in Animal Experiments (Anon [a], 1991)
[4] Finland (Purves, 2000)
[5] Germany (Purves, 2000)
[6] Netherlands (Anon [a], 2004)
[7] NSW Agriculture (Anon [c], 2003)
[8] Swiss Federal Veterinary Office (Anon [d], 2004)
[9] U.K. (Purves, 2000)
[10] USDA – APHIS (Anon [e], 2004)

recognising that the total impact on an animal depends on more than the specific manipulation being carried out, e.g. the competence of those carrying out the manipulation. However, the review also identified areas where greater clarity of purpose and more effective practical implementation could be achieved.

These included:

- Use of more appropriate and balanced descriptors of steps on the scale;
- Ensuring that the scale categories collectively cover the whole range of possible manipulations;
- Ensuring greater accuracy in categorisation by predicting impact on individual animals or groups within an experiment rather than giving one grade to a whole experiment;
- Ensuring greater accuracy of statistics by requiring that grading reported to MAF reflect actual rather than predicted impact;
- Ensuring that category descriptors and examples are seen as guidelines rather than as prescriptive requirements, thus emphasising the importance of judgement when assessing individual cases;
- Determining whether the scale will adequately cover new developments such as those used in biotechnology;
- Ensuring that all factors with the potential to impact on animals are taken into consideration, for example the provision of enrichment, special nursing techniques, and the use of analgesia.

Given the role of the classification system in following statistical trends in the distribution of manipulations from year to year, as well as the need to ease transition to a modified scale, there would be value in ensuring that any changes to the system are compatible with the previous scale.

The influence of terminology – "impact" *vs.* "severity" and "suffering"

The Mellor and Reid (1994) paper discusses potential welfare compromise, ethical cost and the impact of procedures on animals. It outlines a rigorous method, based on functional criteria and informed judgement, to rank the severity of manipulations and assign them to specified categories. This has commonly become known as the "severity scale". In practical terms, on animal ethics application forms and on the statistics forms required to be sent back annually by institutions using animals to MAF, the five grades, each of which is explained in some complexity within the body of the Mellor and Reid (1994) paper, have been

reduced to the following simple definitions – no suffering or virtually no suffering (0), little suffering (A), moderate suffering (B), severe suffering (C) and very severe suffering (X) (Anon (b), 2001).

Given that one of the reasons for reviewing the scale was to assess whether it best meets the needs of all those who use it and the interested public, it was important to determine if misconceptions can arise simply from the naming of the categories. Both "severity" and "suffering" are words with negative connotations. While they may be appropriate when the degree of welfare compromise is high, it is perhaps misleading to label the whole grading system in this way. For instance, in the New Zealand context at least, animals which may be manipulated simply by a change in diet, thus falling into the minimal impact "0" category, are still categorised in terms of "suffering", albeit "little or no suffering". The concept of suffering has a particular meaning when used in the assessment of welfare, including as it does varying degrees and combinations of anxiety, fear, pain and distress. In a wider "lay" context, however, it is an emotive word and one that can have very different meanings to different people. The word "severity" in relation to the grading scale likewise sets a negative context.

The recommendation to replace both "severity" and "suffering" with the more neutral word "impact" is not in any way meant to imply that some animals do not suffer when used for research, testing and teaching. One of the impacts on animals may indeed be suffering, but having the words "severity" and "suffering" to describe the scale is clearly not precise when they are already used as descriptors for categories *within* the scale. Rather, the change is an acknowledgement that while there is always an impact, suffering does not always occur. "Impact" covers all effects that any procedure may have on an animal, while "suffering" does not. Thus the words "suffering" and "severity" may be used within categories of the "impact" scale, but do not define them.

Number of categories

The need to classify the wide range and diverse characteristics of manipulations in animal-based research, teaching and testing into categories has been well established. The present system aims to cover all possible procedures that might be carried out on animals used in research, testing and teaching, but it is suggested here that classification systems should also include manipulations, referred to by Banner et al. (1995), which would not be allowed under any circumstances.

The comparison of scales used internationally showed that most used between three and five categories, and reviews have raised suggestions that overcomplication may result from too many categories (Smith and Jennings, 2004). While New Zealand's current 5-category scale starts at the lowest level, where the manipulations involve no or very little invasiveness or severity, and progresses to the highest such severity rating that would only very rarely, and for the most compelling of reasons, be approved, it is recommended that a sixth category of manipulations be added to include manipulations, which, while quite possible, are of a severity that is unacceptable. It is proposed that this grading be designated Z. Modified thus, the impact scale would then encompass the full range of possible manipulations. This puts the range in true perspective as it indicates that the small number of manipulations that are graded X, while of high impact, are not as high as is possible. Without this extra grading, the impression is conveyed that X manipulations are "as bad as it gets". If this were to be adopted, the annual national score for this sixth category (Z) would therefore always be zero.

Segregation of manipulation categories within an experiment

The Animal Welfare Act 1999 lacks some clarity in how different animals used in a single experiment are categorised. Quite frequently there is a marked difference between groups within an experiment. For example, severe impact groups may be compared with negative controls. This can, and has been, interpreted in two ways. Either the whole experiment is categorised according to the most severe category or groups are segregated and individual groups are categorised separately. In collecting statistics, it is essential to be as accurate as possible, but it is as important to be consistent.

Greater accuracy is obtained if scoring is carried out separately on segregated groups or even individuals. Untreated controls, for example, would not be counted with other groups of treated animals, which would inflate the number of animals receiving high impact treatments. This is particularly important in experiments in which large numbers of animals are compared in groups, but also, for example, where a series of groups of animals may be set up in which the effect of a substance is measured as a dose response. The groups may vary as a graded series from nil to high impact, requiring a measure of judgement on behalf of both researchers and animal ethics committees.

Reporting of actual, rather than predicted, impacts

The prediction of impact is necessarily part of an application for ethics approval prior to animal manipulations taking place. In many cases, particularly those where the researcher has extensive experience with a particular manipulation, those predictions will be accurate. However, this is not always the case. There may be instances where a manipulation results in a lesser or indeed a greater impact than predicted.

It is important in terms of accuracy of the statistics that the grading listed in the annual returns to MAF reflects not the predicted impact but the actual impact. This requires that those manipulating animals keep accurate records of the actual impact during the experimental or testing period for subsequent reporting, a practice that is not universally carried out at present. This practice not only enables greater accuracy of reported statistics, but should also allow more accurate predictions for similar manipulations in subsequent experiments. This is important for both the applicants and the AEC members, for whom such information informs the monitoring process, which is part of their remit.

Category descriptors as guidelines

It is important that the grading system is not seen as inflexible. It is not possible to produce a workable scale that specifically takes account of all possible manipulations in all possible

species. This necessitates the use of various qualitative magnitude terms such as "short term", "rapid", "protracted", etc., where what is meant very much depends on a number of factors, such as species and type of manipulation. This means that any assessment or prediction of impact will require a measure of judgement on the part of both applicant and AEC. It also underlines the importance of regrading manipulations upon completion of the study (as noted above), not only to provide accurate statistics, but also to widen the knowledge base on the effect of such interventions.

Animal manipulations in biotechnology

There is a public perception that the accelerating development of biotechnology has provided completely new challenges with respect to the ethics of using animals in research. As a result there has been an active societal debate on the acceptability of biotechnology, particularly in relation to specific processes such as genetic manipulation.

In the context of this discussion, biotechnology can be defined as the use of techniques such as those of "molecular biology" to effect functional changes at the tissue and/ or whole-animal level. Changes may be temporary or may become a heritable trait.

Leaving aside the wider ethical question of the acceptability of biotechnological applications in general, the concern here is with the more specific question of whether such manipulations are acceptable in animal welfare terms. That is, whether they meet currently appropriate and generally accepted ethical criteria for animal use in research and can be assessed with the expertise currently available on ethics committees as set up in New Zealand.

Examples of such manipulations include:

(i) Manipulating the gamete (ovum, spermatozoon) or embryo;
(ii) Addition to or deletion from the genome;
(iii) The cloning of animals;
(iv) The repopulation of tissues or creation of whole animals with stem cells;
(v) The secretion of proteins originating from other organisms;
(vi) Xenotransplantation.

The following observations on some of these examples can be made:

- At the present time, manipulating the ovum, during cloning for example, does not require ethical approval in New Zealand, because an ovum is not considered to be an "animal" under the Animal Welfare Act 1999. Moreover, the associated whole-animal manipulations are identical to, and standard for, *in vitro* reproductive manipulations, which are now routine and have received AEC approval for some time. However, the potential for such manipulations to affect the phenotype of resultant offspring has led to a review currently being conducted by the National Animal Ethics Advisory Committee of the exclusion of ova manipulation from the need for ethical approval. (See final bullet point below for monitoring of animals with new phenotypes).
- Changing the genome of animals with the use of mutagens has been a process carried out for many years, although now largely replaced by biotechnology. Screening for mutations is already carried out in investigations of unwanted side effects with candidate anti-cancer drugs under AEC approval.
- Stem cell biology has been studied for many years. Altering stem cells prior to transfer is carried out *in vitro* and therefore presents no novel ethical considerations for AECs, which focus on the welfare implications of the introduction of such cells into whole animals.
- The xenotransplantation debate mainly involves animal-to-human transfers. Notwithstanding this, and acknowledging that obtaining pathogen-free supply animals does have special welfare considerations, any associated consequences can be well managed by current AEC procedures that have applied to animal-to-animal transfers conducted for many years.
- Breeding offspring from founder animals may appear to be novel, but the processes have been long carried out in mice in producing inbred lines and recombinant inbred lines. Clearly, the presence of new genes, for example, may produce different phenotypes depending on the gene background and on whether the animal is heterozygous or homozygous with respect to the new gene, which may have negative animal welfare impacts. Again, these can be well managed by current AEC procedures.

In summary, whether animals used in biotechnology are the providers or carriers of manipulated genetic material, the resulting offspring and even the subsequent generations, related animal welfare considerations can provide some new challenges in terms of rate and scale of change. The vast majority of the manipulations involved are variations on current technologies which have provided few problems during AEC assessment. Although extra care will be needed to monitor the welfare of animals having phenotypes with unknown welfare implications, this requirement falls within the current remit of animal ethics committees. This is supported by a paper (de Cock Buning, 2004) which concluded that, as long as the welfare of animals at all stages of a biotechnological procedure is assessed by an animal ethics committee, the wider discussion surrounding these processes should be addressed in a different forum.

Factors impacting on animals in research, teaching and testing

The five domains of compromise, which relate to the nutritional, environmental, health (injury, disease), behavioural and mental needs of animals as suggested by Mellor and Reid (1994), provide an appropriate basis for classification, as they allow an holistic approach to assessing impact – an approach that has been endorsed by other reviews of such categorisations (Orlans, 2000; Smith and Jennings, 2004). The current scale goes through physical impacts first and then addresses the mental state of the animal. However, it is now suggested that a provisional score with respect to mental state be established first as the ultimate measure of impact, with the other four as contributors to that ultimate measure being double-checked to ensure no

factor has been missed, nor the impact with regard to mental state over or underestimated.

Alongside and contributing to the "domain" assessment, both intrinsic (animal) and extrinsic (environmental, procedural) factors that will influence the impact on the animal need to be taken into account. These include:

Animal factors

- Species/type/breed differences – marked variation between and within species can occur in terms of, for example, responses to human presence, fear responses and temperament. A detailed understanding of unique and common features of the biology and behaviour of each species studied, including life stage differences, is essential in assessing welfare impact. Moreover, an adequate application of the scale may necessitate consultation with others with greater expertise in the species of interest.
- Size – the impact of factors such as needle gauge, injection volumes and extent and complexity of surgical manipulations relative to the size of the manipulated animal needs to be accommodated.

Environment

- Provision for social and behavioural needs, including the enrichment of environments, should be taken into account in the categorisation of procedures. Likewise, the influence of physical environmental features, including temperature, light, sound, air quality, vibration, space, comfort of ground/floor surfaces, safety and other aspects need to be assessed. Also, the impact of animal attendants and investigators who are part of the animals' environment should be considered.

Procedures

- Types of manipulation – the welfare impact of each manipulation relates to both the magnitude of negative effects and their duration, as well as to the interval between successive manipulation, e.g. for recovery. The five domains of potential welfare compromise (Mellor and Reid, 1994) provide a comprehensive way of assessing impact. Any guidelines relating to specific procedures should be indicative, not prescriptive, requiring the exercise of good judgement when being applied.

The final grading will depend on other factors as well, including:

- Operator skill – there may be a need to grade skilled operators and trainees differently, for example. This emphasises the need for some flexibility in grading procedures.
- Harm minimisation – the extent of application of the 3Rs to proposed manipulations will affect the final grading.

Summary

1. The basis of the system in the Mellor and Reid (1994) paper remains appropriate.
2. The name of the categorisation system should be the "impact scale".
3. The current 5-point system should be enlarged by the addition of a sixth category which includes procedures that must not be carried out under any circumstances. To minimise confusion in the transition from one scale to another, it is suggested that the original categories retain their symbols of 0, A, B, C and X, while the sixth should be labelled Z.
4. To ensure greater accuracy where individuals or groups within an experiment are likely to experience significantly different impacts, they should be graded on an individual or group basis, rather than a whole experiment basis as originally suggested by Mellor and Reid (1994).
5. Animal manipulations in biotechnology are adequately covered by the present system and the addition of special categories is not necessary.
6. An exhaustive list of manipulations with recommended gradings is not advisable, both because it will inevitably be incomplete and because it tends to be viewed in a rigid manner. Implementation of all the recommended features will require judgement to be exercised by scientists and AEC members alike.
7. A requirement should be made for the predicted impacts of manipulations to be reconsidered at the end of each study to ensure that the actual impacts, as assessed by researchers, animal-care staff, animal welfare officers and AECs, are supplied in the annual statistical returns to MAF.

References

Anon (a) (1991). Categories of invasiveness in animal experiments. Canadian Council on Animal Care. http://www.ccac.ca/english/gui_pol/policies/CATEG.HTM Accessed 14/11/04.

Anon (c) (2003). Form L Animal Use Statistics, New South Wales Agriculture. http://www.animalethics.org.au/reader/arrp-forms/form-l2003.doc?MIvalObj=19990&doctype=document&MItypeObj=application/msword&ext=.doc Accessed 11/01/05.

Anon (d) (2004). Classification prospective des experiences sur animaux selon leur degre de gravite (categories de contrainte), Swiss Federal Veterinary Office. http://www.bvet.admin.ch/tierschutz/00237/00833/index.html?lang=fr&download=00423_fr.pdf Accessed 11/01/05.

Anon (a) (2004). Taking animal welfare seriously: minimizing pain and distress in research animals. Humane Society of the United States. http://www.hsus.org.ace/Article_Printer_Friendly?Content_ID=11425 Accessed 12/07/04.

Anon (b) (2001). *Animal use statistics*. Wellington, Ministry of Agriculture and Forestry.

Banner, M., Bulfield, G., Clark S. et al. (1995). *Report of the Committee to Consider the Ethical Implications of Emerging Technologies in the Breeding of Farm Animals*. London: Ministry of Agriculture, Fisheries and Food (MAFF).

Bayvel, A. C. D. (2004). Animal Use Statistics: The New Zealand and Australian Experience and Perspective. *ATLA 32 (Supplement 1)*, 287-292.

de Cock Buning, T. (2004). Practical difficulties in balancing harms and benefots in the modern use of laboratory animals: biotechnology. *ATLA 32 (Supplement 1)*, 459-463.

Mellor, D. and Reid, C. S. W. (1994). Concepts of animal well-being and predicting the impact of procedures on experimental animals. In *Improving the Well-being of Animals in the Research Environment* (pp 3-18). Sydney, ANZCCART.

Orlans, F. B. (2000). Practical applications of animal harm scales: international perspectives. In M. Balls et al. (eds), *Progress in the Reduction, Refinement and Replacement of Animal Experimentation* (1049-1056). Bologna: Elsevier Science.

Orlans, F. B. (2000). Public policies on assessing and reporting degrees of animal harm: international perspectives. In M. Balls et al. (eds), *Progress in the Reduction, Refinement and Replacement of Animal Experimentation* (1075-1082). Bologna: Elsevier Science.

Purves, K. E. (2000). Global harmonisation of pain and distress classification systems: current analysis and opportunity for change. In M. Balls et al. (eds), *Progress in the Reduction, Refinement and Replacement of Animal Experimentation* (1057-1067). Bologna: Elsevier Science.

Smith, J. and Jennings, M. (2004). Categorising the severity of scientific procedures on animals The Boyd Group and the RSPCA.

Acknowledgements

This review was commissioned and funded by the New Zealand Ministry of Agriculture and Forestry. The authors wish to thank the organisers of the 5th World Congress for the opportunity to present this review, and Dr John Schofield of Otago University, New Zealand, and Professor David Morton of Birmingham University, United Kingdom, for their helpful comments.

Correspondence to

Virginia M. Williams
School of Natural Sciences
Unitec
PB 12025
Pt Chevalier
Auckland, New Zealand
e-mail: vwilliams@xtra.co.nz

A Wider Interpretation of the Three Rs Model

Flavia Zucco
Istituto di Neurobiologia e Medicina Molecolare, CNR, Roma, Italia

Summary
The project Anim.Al.See, which aims to update the 3Rs model, has explored in detail, both from the philosophical and the scientific point of view, the issue of alternatives to animal experimentation. Concepts and language involved in specific case studies related to each R have been analysed. On this basis, new definitions, among others of replacement, reduction, and refinement, have been coined in order to implement a more advanced version of the Russell and Burch model. This has been achieved by widening the framework of reference and thus offering a wider range of approaches.

Keywords: alternatives, three Rs, scientific procedure, sentiency

"Indulge your passion for science, but let your science be human, and such as may have direct reference to action and society."
D. Hume (1711-1776)
Enquiry concerning human understanding. Introduction, 9.

Introduction

About 50 years have gone by since the publication of "The Principles of Humane Experimental Technique" in 1959 by the British scientists Russell and Burch. In this book they challenged the scientific community to take the proposal to improve the treatment of animals used in scientific procedures seriously, and to this purpose they introduced the concept of the 3Rs: replacement, reduction and refinement.

Even if this book has produced efforts by some scientists to improve animal experiments and to search for alternatives (mainly defined as *in vitro* methods), after so many years the debate between proponents and opponents of animal use in scientific procedures seems to have stalemated. A number of aspects, most of them existing since the 1980's, have led to this standby.

The most relevant aspect has been the slow advancement of *in vitro* models in the area of toxicology and pharmacology. While in all areas of biomedicine, *in vitro* cellular models were widely and successfully adopted, toxicology has remained almost impermeable to their adoption for many years, due also to the strict limitations of international regulations requiring testing of chemical substances on animal models before they can be introduced on the market (Council Directive 67/548/CEE and subsequent amendments, 1967). However, it should also be admitted that the prevailing culture among scientists in the field was that of "*in vivo veritas*". This statement was surprisingly contradictory, because, at the same time, it was clear also to toxicologists that basic toxicity mechanisms can be better investigated in isolated systems, which avoid the complexity of the whole organism. Thus, toxicology appeared a very conservative discipline, poorly amenable to novelty (Paganuzzi-Stammati et al., 1981; Zucco et al., 2004).

The second aspect was that the Russell and Burch model (3Rs) was subsequently defined as the "3Rs alternative model" by Smyth, and thus the focus has been mainly on replacement. Indeed, in the original Russell and Burch proposal non-animal approaches were only one part of the story (replacement). Reduction and refinement, which should ameliorate the situation of animals in the labs, were also addressed.

The third point was that Smyth's definition (1978) of the 3Rs as "alternatives" has created confusion, giving rise to an endless discussion about whether they should be called alternatives or (better) complementary or adjunct strategies. When Smyth referred to the 3Rs as alternatives, he was probably not simply referring to possible options but, according to the Webster Dictionary, to "an option existing or functioning outside the established cultural, social or economic system". However this choice of word has been misleading the debate, which became biased by exceeding expectations.

The final aspect was that the 3Rs model mainly addressed the scientific community, so that it was of no help in the public debate on animal experimentation, where pro and contra stakeholders, including the society at large, were confronting each other with opposite positions without any progress.

Moreover, it should be taken into account that from that time, much scientific advancement has been reached, both in terms of knowledge as well as of technology, so that the 3Rs can be discussed with a more solid background than before, but also with more challenges to be faced (Schuppli et al.,2004). In the meantime, the culture of the society has changed, and those changes regard science as well as the position of humankind in the natural environment and its relationship with all living species.

For all the above-mentioned reasons, the issue of animal experiments and their possible alternatives needs to be investigated more deeply from the theoretical and ethical point of view in connection with intrinsic aspects of scientific research and the present social culture. Theoretical and ethical aspects, indeed, deserve to receive as much study and elaboration as practical

and scientific aspects in terms of research into, and implementation of, alternatives.

In 1999, the project "Alternative methods in animal experimentation: evaluating scientific, ethical and social issues in the 3Rs context" (Anim.Al.See: *Anim*al *Al*ternative: *S*cientific and *E*thical *E*valuation) was proposed to the European Commission under the Bio-ethics activity of the Quality of Life Programme (FP 5).

The Anim.Al.See project

The aim of Anim.Al.See was to assess scientific and philosophical aspects together to update language and concepts. It was clear that too many years had passed since the 3Rs model had been proposed and that too many changes had happened in many different fields related to the issue of animal experiments. The major issue was that technological progress had not been followed by adequate philosophical and ethical analysis. The need for a more advanced ethical inquiry was justified by impressive changes in the societal culture. For this reason the project was carried out by both scientists and philosophers.

The Science

Science is no longer seen at the core of the progress: different tools are needed to cope with the technological advancement in order to translate it into progress in civilisation.

The criticism to which science has been submitted in the last decades, not only by the public but also by experts and scientists themselves, has been relevant in revealing the changes that science has undergone in the most recent decades: no longer authoritative, often raising uncertainty by controversial viewpoints, insufficiently transparent, intrinsically bound to technology and thus to market. The practice of science should thus be reconsidered, and since technology is continuously offering new opportunities but also opening new questions of a scientific and ethical nature, the scientific world is invested with a wider responsibility (Kelley, 1998; Ziman, 2000; Novotny et al., 2001).

The Society

The general societal culture is moving from the historical animalism and anthropocentrism to a global concept of a more equilibrated relationship between humankind and the natural world. The new relevant societal subjects are often the movements and NGO, while the national states are loosing their identity. We are thus facing a transition in approaching and solving problems, searching for alternatives in science and technology. The reference to values and personal, as well as collective, responsibilities expresses the need to direct science towards different ways and goals. One of these is the change from the historical view of discontinuity between animals and humans (rational beings) to a more holistic view of life forms, suggesting a sort of reconciliation between the human species and the natural system.

Recent research has shown that animals are closer to us than thought before, both from genetic and physiological points of view. At the same time, differences between animals and humans have also been highlighted. This brings us to the problem of considering them worthy of moral consideration, and less reliable as models for humans. Our analysis should specifically take into account the new moral concern for animals, which not only rejects their instrumental use, but points at their relationships with humans, based on ethical concerns and our new knowledge on animal sentiency.

The semantics

The issue of bridging the gap between the humanistic and scientific cultures seemed a relevant aspect to be taken into account in this research area (see also Fraser, 1999; Porter, 1992). This debate indeed dates back to the 1950's, particularly to a book published by C. P. Snow in 1959. The essence of this debate is that science can no longer be considered the only engine of progress: it must be accompanied by humanistic approaches to readdress the question of progress in civilisation. Technological progress (application of science) needs to be combined with responsible ethical analysis, grounded in humanistic approaches (sociology, history, psychology, philosophy).

It was thus clear that the work should also focus on the revision of the semantics, i.e. the language, definitions and the concepts behind them. This is indeed a current topic in science at large. Scientific language is becoming less and less appropriate to address new concepts and theories emerging from the advancement of knowledge. Several disciplines have been dealing with this aspect, which not only gives rise to problems in communication with society, contributing to a lack of transparency, but also affects scientific disciplines (Appleby, 1999; Bensaude-Vincent, 2001; Kiessling, 2001; Sloviter, 2002; Wallace, 2002; Wilczek, 2001). It is increasingly admitted that the analysis of theoretical concepts (and assumptions of the framework of reference) and of the language used is indispensable to progress in specific fields of science. This is certainly the case for the 3Rs and their definition as alternatives.

Confrontation with the above-mentioned aspects has led to a broadening of the meaning of the 3Rs model and to provision of new definitions of replacement, reduction and refinement, as well as of alternatives. Thus, while the model still mainly addresses the scientific community, it now also considers the issue of responsibility towards other subjects, such as public and private institutions, regulatory boards and the public at large, and includes ethical evaluation as a factor with the same weight as scientific evaluation in areas where multiple interests are involved.

Methodology

The project was developed in three phases

1. Phase One

Each partner obtained an updated background of current practices in the 3Rs as alternatives to animal experiments. The preliminary step was to define key concepts of the different aspects involved.

A questionnaire (in electronic format) was then formulated, assisted by an expert in social surveys, in an endeavour to

enquire into the conceptual background. The mailing list was carefully drawn up in order to cover the main experts in the different fields, i.e. scientists, regulatory boards, industry, consumers, animal associations and ethicists, but, at the same time, was targeted to avoid dispersion.

The responses were analysed and the results compared to the definition of the key concepts on which the questionnaire had been formulated (Pollo et al., 2004).

2. Phase Two

Three different interdisciplinary groups challenged the outcomes of the first phase against case studies chosen as relevant for each "R". Thus, the extent to which the framework, developed during Phase One, reflected the reality of the problems raised by practical approaches to alternatives to animal experiments was verified. Two case studies were analysed as being relevant to the concept refinement: experimental protocols involving non-human primates, and housing of non-human primates in biomedical research. In both cases, a philosopher analysed the appropriateness of the criteria used to define the ethical aspects of the research when these criteria were applied to the welfare of non-human primates in biomedical laboratories, whereas primatologists focussed on the possibility of applying non-invasive procedures in experimental protocols and environmental enrichment in housing conditions.

From this analysis the following definition of refinement was elaborated:

Refinement refers to any approach which avoids or minimises the actual or potential pain, distress and other adverse effects suffered at any time during the life of the animals involved, and which enhances their well-being.

Regarding reduction, the "single-dose" test alternatives to the multi-dilution test in vaccine quality control and the case of telemetry were analysed. These examples are of interest for various reasons: traditional testing uses large numbers of animals, a high level of distress is involved, and the tests are required within a regulatory framework. Telemetry is interesting due to its recent widespread use as a possible means of reducing the numbers of animals used, but it can conflict with refinement.

The proposed definition of reduction is:

Reduction refers to any approach in scientific research, product testing or education that leads, directly or indirectly, to a decrease in the number of animals used.

The case study on replacement dealt with issues of *in vitro* methods as alternatives in cosmetic testing. This area has a high impact on public opinion and has produced a series of amendments to the legislation, mostly concerning postponing the European ban on animal testing for cosmetics. An evaluation has been performed of the interplay between the different aspects underlying new tests and their development, and how they comply with the existing conceptual framework. The relative importance of each aspect has been assessed in relation to the practicability and feasibility of possible innovations, as well as the ethical issues involved in the so-called alternatives. The following definition has been proposed:

Absolute replacement refers to any approach in scientific research, product testing, and other technical procedures in which no animals are used.

The *ex vivo* aspects of cell cultures – serum provision and primary cell cultures – are considered to fall under reduction and refinement practices.

3. Phase Three

In Phase Three, the work was concentrated on producing an updated version of the "3Rs" model, according to the outcome of the investigations of the case studies in relation to the current cultural context.

A short executive summary of the unified report and the drafted recommendations have been submitted to selected experts invited to a final conference scheduled shortly before the termination of the project, in order to obtain the opinion of representatives of the different stake holders. The outcome of the conference has been integrated in the final edition of the project report.

Beside the redefinition of the 3Rs, new definitions of "alternatives" have been elaborated: A narrow definition referring to the scientific practice is as follows:

An alternative to an animal experiment is any procedure – i.e., any method or technique, proposal or approach – that is meant to replace a particular science-based procedure that may harm the interests of animals, to reduce the numbers of animals required, or to refine the procedure in such a way that the welfare of the animals in the procedure itself or in its context, is optimised.

A wider definition offering wider possibilities of solution, also external to the scientific context, is the following:

An alternative to animal experiments is any ethical, cultural, political choice, which is meant to abolish animal scientific procedure, on the basis of a responsible societal agreement supported by the awareness of the implication concerning science, policy etc.

Moreover we proposed to adopt "scientific procedure" instead of "experiment" and defined it as follows:

A combination of one or more acts carried out on an animal for an experimental or other scientific purpose, and which may cause that animal pain, suffering, distress or lasting harm.

Finally we provided, following all the considerations emerging from the work done on animal care and welfare, a definition of "sentient animals":

Sentient, non-human vertebrates, and other species of animals that have interests and capacities that are comparable to those of vertebrates.

Conclusions

In summary, the main conclusions emphasise the need to:

- Acknowledge that the ethical dimension to animal experiments is as important as the scientific dimension.
- Clarify concepts underpinning current terminology concerning alternatives and animal experiments, methods, procedures, etc., to recognise the role played by semantics in this debate.
- Adopt a narrow and a wider definition of alternatives related to the socio-political context.
- Reinforce the "3Rs" model as a sound scientific approach to improve animal experimentation.

- Adopt the updated definitions of the principles of Replacement, Reduction and Refinement.
- Promote the application of alternative procedures to animal experiments, according to this new framework of reference.

Beside the specific aspects mentioned above, a more general feature should be pointed out: the theoretical approach and the practical methodology adopted in this project may be useful in other areas of conflict between science and society.

The most important aspects that can be transferred to other areas are:

- Honest, exhaustive and interdisciplinary analysis of the issue;
- A precise and commonly accepted use of language;
- Making scientific aspects more accessible, reducing technicalities and making explicit embedded concepts and values;
- The importance of an appropriate philosophical background to understand the relevance of ethical evaluation in technological implementation.

For more information (documents, recommendations, etc.) on the project Anim.Al.See, consult the website www.inemm.cnr.it/animalsee.

References

Appleby, M. C. (1999). Tower of Babel: variation in ethical approaches, concepts of welfare and attitudes to genetic manipulation. *Animal Welfare 8*, 381-390.

Bensaude-Vincent, B. (2001). Chemical analysis – Language reform played an integral role in the development of a discipline. *Nature 410*, 415.

Council Directive 67/548/EEC of 27 June 1967, (16-8-1967). On the approximation of laws, regulations and administrative provisions relating to the classification, packaging and labeling of dangerous substances. *Official Journal of the European Communities 196*, 1-98.

Fraser, D. (1999). Animal ethics and animal welfare science: bridging the two cultures. *Applied Animal Behaviour Science 65*, 171-189.

Kelly, K. (1998). The third culture. *Science 272*, 992-993.

Kiessling, A. A. (2001). In the stem-cell debate, new concepts need new words. *Nature 413*, 453.

Novotny, H., Scott, P. and Gibbons, M. (2001). *Re-thinking science*. Blackwell Publishers Inc. USA.

Paganuzzi Stammati, A., Silano, V. and Zucco, F. (1981). Toxicology investigations with cell culture systems. *Toxicology 20*, 91-153.

Pollo, S., Vitale, A., Gayle, V. and Zucco, F. (2004). The "3Rs" model and the concept of alternatives in animal experimentation: a questionnaire survey. *Lab. Animal. 33*, 47-53.

Porter, D. G. (1992). Ethical scores for animal experiments. *Nature 356*, 101-102.

Russell, W. M. S. and Burch, R. L. (1959). *The Principles of Human Experimental Techniques*. London: Methuen.

Schuppli, C. A., Fraser, D. and McDonald, M. (2004). Expanding the Three Rs to meet new challenger in humane animal experimentation. *ATLA 32*, 525-532.

Sloviter, R. S. (2002). Apoptosis: a guide for the perplexed. *Trends in Pharmacological Sciences 33*, 19-24.

Smyth, D. H. (1978). *Alternatives to animal experiments*. London: Scolar Press.

Snow, C. P. (1959). *Two cultures and the scientific revolution*. Cambridge UK: Cambridge University Press.

Wallace, A. (2002). The emerging conceptual framework of evolutionary developmental biology. *Nature 415*, 757.

Wilczek, F. (2001). When words fail – Scientists have to struggle with words that don't fit reality. *Nature 410*, 149.

Ziman, J. (2000). *The real science. What it is, and what it means*. Cambridge UK: Cambridge University Press.

Zucco, F., De Angelis, I., Testai, E. and Stammati, A .(2004). Toxicology investigations with cell culture systems: 20 years after. *Toxicology in Vitro 17*, 153-163.

Acknowledgments

The Anim.Al.See project was financed by the European Commission, Contract: QLRT-2001-00028. The author is indebted to all the partners who have contributed to the project: Caterina Botti of INEMM (Istituto di Neurobiologia e Medicina Molecolare, CNR, IT); Annalaura Stammati, Augusto Vitale and Simone Pollo of ISS (Istituto Superiore di Sanità, IT); Hannah M. Buchanan-Smith and Anita Rennie of UOS (University of Stirling, UK); Coenraad Hendriksen and Jasmijn de Boo of NVI (Netherlands Vaccine Institute, NL).

We are all specially grateful to Jan Vorstenbosch, Institute of Ethics, EINL (University of Utrecht, NL), who was able to produce an optimal coherent elaboration of our varied ideas.

Correspondence to

Dr. Flavia Zucco
Coordinator of the EU Research Project
Anim.Al.See, QLRT-2001-00028
Istituto di Neurobiologia e Medicina Molecolare, CNR
Via del Fosso di Fiorano 64
00143 Roma, Italy
e-mail: f.zucco@inmm.cnr.it

Session 3.6
Establishing the 3Rs Principle in Japan

Education in Alternatives to Animal Experimentation

Tsutomu Miki Kurosawa
The Institute of Experimental Animal Sciences, Osaka University Medical School, Japan

Summary
This article describes various aspects of 3Rs education. To be able to understand and discuss 3Rs education, a certain background knowledge of animal experimentation, in particular animal experimentation as an alternative to human experimentation, is essential. The realities of animal experimentation and medical and veterinary needs are discussed. The initiation and consequence of medical research is explained using examples of the author's research activities.

Keywords: alternatives to animals, animals in education, three Rs education

Introduction

Education in alternatives to animal experimentation has spread to all parts of the world including Asian countries (Kurosawa, 2004a). We should aim to educate all types of audience, from school children to professors and the general public, on the concepts and importance of the 3Rs (Kurosawa, 2004b; Cervinka and Kurosawa, 2004). Naturally, the approach must be specific to the target group: the approach to inform the general public must differ from the approach to postgraduate biomedical students.

The first step of 3Rs education should be education on the reality of animal experimentation. If there were no animal experiments, we would not be here to discuss them at the 5th World Congress. We need to understand why animal experiments are still being carried out. And we need to realise and express why we believe animal welfare and the 3Rs are important.

Animal experimentation

Animal experimentation is the alternative to human experimentation. Animal experiments are carried out when the same experiments on human beings cannot be ethically or morally justified. Particularly during the Second World War many such experiments were carried out on citizens and captured soldiers. As humans are the most relevant model for human physiology, biomedical progress was truly significant during that time. When this practice became public after the war ended, strict prohibitions were set on experiments on humans, particularly in the biomedical field. Still, the victorious countries secretly moved the records of human experiments carried out during the Second World War to their home countries.

Examples of medical progress achieved by animal experimentation

One of the most significant medical milestones was the first successful heart transplantation. The main function of the heart is to circulate the blood to various parts of the body. This fact was established some 300 years ago by pioneering biologists performing animal experiments. At that time, scientists and doctors tended to use the animals that were easily available and consequently they used many pet animals such as dogs and cats. The mechanics required for heart transplantation, i.e. sewing up the vessels, were already clear very early, but the most difficult challenge to successful heart transplantation was the immunological rejection of the donor heart. Thousands of mice and rats and other laboratory animals were required to solve the immunological rejection of organs, and we finally discovered drugs that introduce immunotolerance. After the discovery of these drugs, not only heart transplantation but also many other organ transplantations became possible and many patients have since then been permanently cured by organ transplantations. Unfortunately, donor hearts are in short supply and scientists are looking for novel ways to cure patients. Some scientists are try-

ing to produce human-like hearts in pigs with transgenic technology and they tend to test the safety of the transgenic pig hearts for humans on primates due to the similarity of immunological function in humans and primates.

Other examples of medical progress were achieved by chance discoveries. The very initial neurological studies using the seaslug might have been intended as a purely biological activity instead of medical research. However, the investigation of the neurological function of the seaslug led to the discovery of the scientific principle behind the cause of Parkinson's disease in human beings: the shortage of dopamine in certain neurons. The investigators were awarded the Nobel Prize in Physiology or Medicine in 2000.

Why do we continue to perform animal experiments?

Because there are still many incurable diseases in the world.

The number of kidney disease patients is increasing very rapidly and in the last year, the number of chronic renal failure patients was over 250,000 in Japan, but no cure-oriented therapeutics are available at present. Personal as well as scientific interests led the author to seek such treatment methods. An animal model of this disease appeared to be the most promising approach. The discovery of a strain of naturally occurring nephritic mice led to the establishment of a chronic renal failure mouse strain. Testing of various possible therapeutics in this model led to the chance finding that HGF (hepatocyte growth factor) could cure chronic renal failure. Application of HGF led to a dramatic recovery of the damaged kidneys (Mizuno et al., 1998). Further animal experiments are now necessary to find a way to apply HGF to chronic renal patients using gene-therapy. Neither line of investigation could have been followed without animal experiments.

Medical progress achieved with animal experiments not only benefits man, many discoveries can be translated to other animals and thus to veterinary medicine. Examples are the use of anaesthetics to enable surgery on animals and the use of medical devices such as X-ray machines for the diagnosis of disease in animals.

Understanding the importance of the 3Rs

Although medical discoveries resulting from animal experimentation contribute to human and animal health, these experiments require the use and often the death of animals. However, apart from having a natural empathy for animals, which should not be suppressed or ignored in the scientific communities, humans have the moral obligation to protect animals. These are the reasons why we are seeking the 3Rs to animal experiments. We need to balance human welfare and animal welfare. This concept is incorporated in the ISO 10993 standards.

ISO 10993 part 2 states, "The protection of humans is the primary goal of the ISO 10993 series of standards. A second equally important goal is to ensure animal welfare and to minimise the number and exposure of the laboratory animals" and that the standard was "developed to ensure the welfare of animals used in biological evaluation testing."

How can we transfer this information to different target groups?

School children should be taught both about the value of animal experiments and the medical progress achieved with them as well as about animal welfare and the 3Rs. These subjects should teach children ethical considerations.

University students must be made aware that many alternatives to animal experimentation are available and that they can contribute to their implementation. A list of some of these can be found on the web site of InterNICHE (http://www.interniche.org/). Apart from the theory, these alternative methods should be integrated into the practical courses as alternatives the traditional animals experiments, e.g. models and simulators, film and video, multimedia computer simulation, student self-experimentation, ethically-sourced animal cadavers, clinical practice, *in vitro* labs.

Of course, both the education of school children and university students on alternative methods can only be achieved by educating medical, veterinary and biology post-graduates, researchers and professors, who are or will become their teachers. This can be achieved by exposing them to the progress of alternative methods, e.g. by inviting professional alternative scientists for lectures, by providing alternative scientific journals and other reference materials, by exposing them to fair public opinion on animal welfare and by organising alternative academic meetings such as WC5.

However, the most important party to be educated is the public citizen as a tax payer. The average citizen neither knows of the reality of animal experimentation nor of our efforts for the 3Rs, although these are generally paid for with tax money. Therefore, we have an obligation to inform the citizen on our progress and our goals.

References

Cervinka, M. and Kurosawa, T. (2004). Report of the workshop on incorporating animal alternatives in training programmes: Specific activities in various countries. *ATLA 32, Suppl. 1*, 591-592.

Kurosawa, T. (2004a). Activities in Japan and other Asian countries & summary of the satellite session. *ATLA 32, Suppl. 1*, 593.

Kurosawa, T. (2004b). JSAAE activities in alternatives to animal experimentation for undergraduate schools: Forum for citizens in primary and secondary schools. *ATLA 32, Suppl. 1*, 509-510.

Mizuno, S., Kurosawa, T., Matsumoto, K. et al. (1998). Hapatocyte growth factor prevents renal fibrosis and dysfunction in a mouse model of chronic renal disease. *J. Clin. Invest. 101*, 1827-1834.

Correspondence to

Tsutomu Miki Kurosawa
DVM, M. Phil., Ph. D., DVCS, DJCLAM
The Institute of Experimental Animal Sciences
Osaka University Medical School
2-2, Yamadaoka, suita-shi
Osaka
Japan 565-0871
e-mail: kurosawa@iexas.med.osaka-u.ac.jp

Theme 4
Information systems and databases

Session 4.1
3Rs database and services – developments worldwide

ECVAM's Database Service on Alternative Methods (DB-ALM) – Online

Annett J. Roi and Manuela Flego
ECVAM, Institute for Health & Consumer Protection, Joint Research Centre, European Commission, Ispra (VA), Italy

Summary
The Db-ALM of ECVAM, requested by the European Commission and Parliament, provides ready-to-use information on various aspects of animal alternatives on evaluated data sheets. The first Internet version has been available since 2001 (http://ecvam-sis.jrc.it). In 2006, access shall be provided, in addition, to the INVITTOX protocol collection, to method-summary descriptions, test results, details on validation studies, and to a new data sector on (Q)SARs. The Db-ALM information content is created in collaboration with external experts or originates from validation studies. The current Internet version can refer to 4667 registered users from 67 countries with a 26 % increase over the past 12 months. The progress of Db-ALM is monitored by ECVAM's Scientific Advisory Committee.

Keywords: European policies, ECVAM, factual databases, animal use alternatives, toxicity tests, in vitro, QSAR

Rationale

The establishment of DB-ALM (formerly known as the "SIS" – Scientific Information Service) is one of the principal objectives of the European Centre for the Validation of Alternative Methods (ECVAM), created by the European Commission in response to a requirement of Directive 86/609/EEC (1986) on the protection of animals used for experimental and other scientific purposes. The DB-ALM shall provide information on the use of animal alternatives in biomedical sciences and serve Commission Services, National Authorities, Industry, the Animal Welfare Movement, as well as scientists.

Information content and supplementary activities

The DB-ALM is characterised by ready-to-use information provided as evaluated data sheets. It covers various aspects of animal alternatives at any stage of development or validation (Janusch and Balls, 1999; Roi and Burke, 2004). A recent initiative led to the extension of DB-ALM to include also information on computer modelling, such as (Q)SARs, in addition to the *in vitro* techniques. The DB-ALM focus on toxicity assessments and a short *résumé* of the main data files, together with an overview of the current information content referring to 21 different topics, are included hereafter:

In Vitro techniques

Methods

The DB-ALM provides full method-descriptions, based on extensive literature reviews with information on the topic area, rationale for the method development and its applications, basic procedure, type of tested materials including the state of development, validation or regulatory acceptance. To address various user profiles, which turned out to be very heterogeneous, the method descriptions are provided at various levels of information content, either as summary record or as detailed protocols (*INVITTOX* protocol collection (Janusch et al., 1997; Ungar, 1993) for their performance in laboratories.

Tests

Tests are described with their experimental details and results carried out with methods/test systems included in DB-ALM.

Validation studies

One of the main tasks of ECVAM is to assess the scientific validity of tests for specific purposes through the co-ordination of complex international validation studies. The DB-ALM provides details on these studies, including summary descriptions of the study, the test protocols and results, and also indicates the participating organisations and lists the background documentation.

Computational modelling – (Q)SARs

A new data sector on (Q)SARs is under development and will provide information on evaluated (Q)SAR models as summary records, together with the training and test set of chemicals, and, where available, summary descriptions of validation exercises.

Supplementary activities

Based on the competence and experience gained in the specific field, the DB-ALM staff is responsible for two additional activities that are briefly outlined hereafter. Both are supported by either the available database tools or by the data content of DB-ALM itself and together make DB-ALM a comprehensive information service on animal alternatives of ECVAM.

The ECVAM website

The ECVAM website, established in 2002, provides details on in-house and international collaborative activities of the entire validation centre. An overview of all validation activities is given, and all validated methods are reported together with a list of the main publications of ECVAM. The website is an interactive service with various download facilities and now also provides support to the various international working groups of ECVAM. An easy-to-use registration facility is available for those who want to be updated regularly on news from ECVAM.

The Thesaurus

A pilot project has been undertaken and concluded by the international Task Force on Alternatives Databases of ECVAM to identify, based on a selected set of publications, the terminology most commonly used by scientists active in animal alternatives topic area. This also in reply to the recommendations outlined in the ECVAM workshop on Alternatives Databases (Janusch et al., 1997; Grune et al., 2004). The availability of this thesaurus, first as an open source list, for practical applications by the end-user, is linked to the release on the Internet of the new DB-ALM version, which is foreseen for the end of 2005/start of 2006.

Tab. 1: DB-ALM Information content "*In vitro* techniques" (August 2005)

Data file	Data sheet number
Method Summary Descriptions	65
INVITTOX protocols	126
Evaluation/Validation Studies	32
Test Results	4219
Bibliographic References	3781

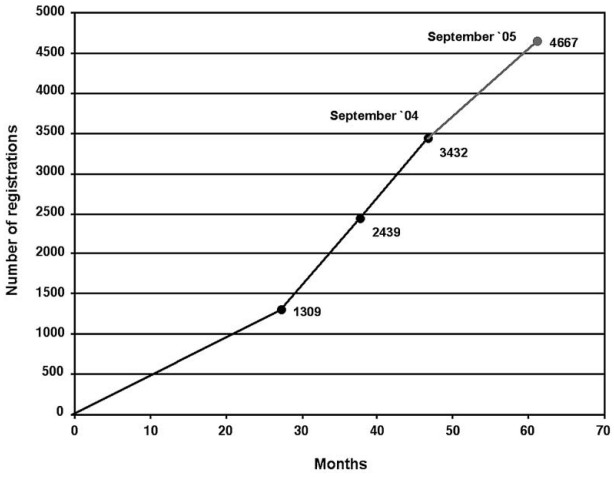


Fig. 1: Registration Tendency (September 2005)

Origin of information

The proper running of the DB-ALM relies on two main "pillars": Compilation of data sheets (creation and updating of the database contents) and Informatics (software developments and hardware requirements for data management and access).

The DB-ALM operates in collaboration with external experts engaged by study contracts and by bilateral contacts between the DB-ALM staff and scientists (particularly for the definition and updating of the *INVITTOX* protocols). It also incorporates the information provided by formal validation studies (Janusch and Balls, 1999).

The priority topics for DB-ALM are defined on the basis of a proposal put forward by the co-ordinator of the DB-ALM at the beginning of each year. The progress of the DB-ALM is monitored by the ECVAM Scientific Advisory Committee (ESAC) and DB-ALM Advisory Board.

Impact analysis of DB-ALM

The user profiles turned out to be very heterogeneous and a statistical impact analysis of DB-ALM in September 2005 revealed that DB-ALM can now refer to 4667 registered users from 67 countries, with a 26% increase in registrations during the past 12 months (fig.1). The current DB-ALM went online first in 2001. Major customers are the USA (22%), UK (15%), India (13%), Germany and Italy (7% each), and France and Spain (6% each). User profiles include: Academia (45%),

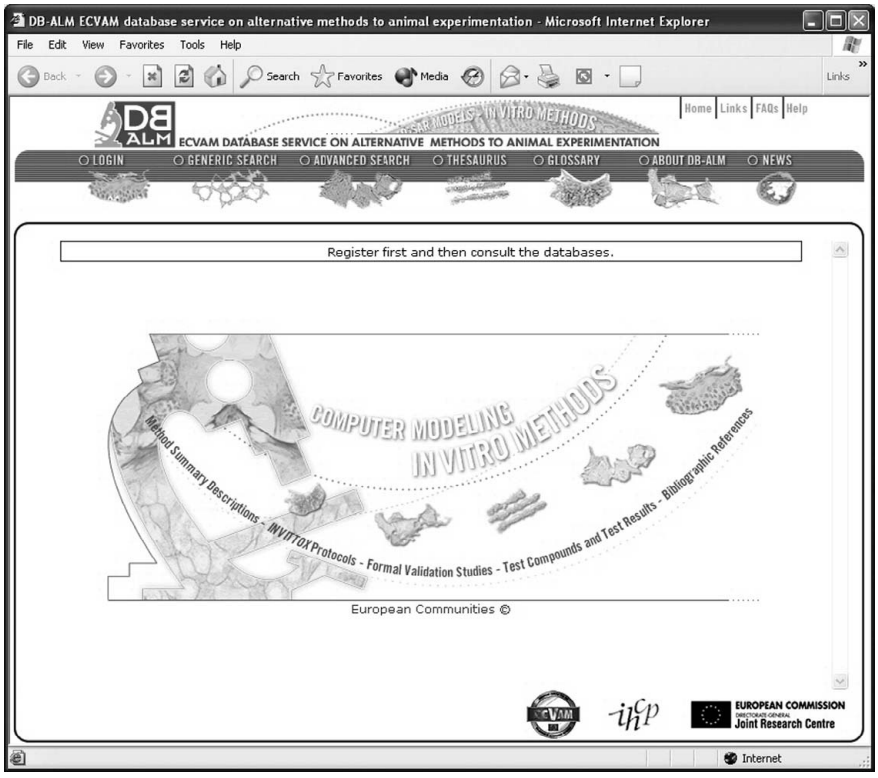


Fig. 2: New Homepage DB-ALM http://ecvam-dbalm.cec.eu.int (valid in 2006)

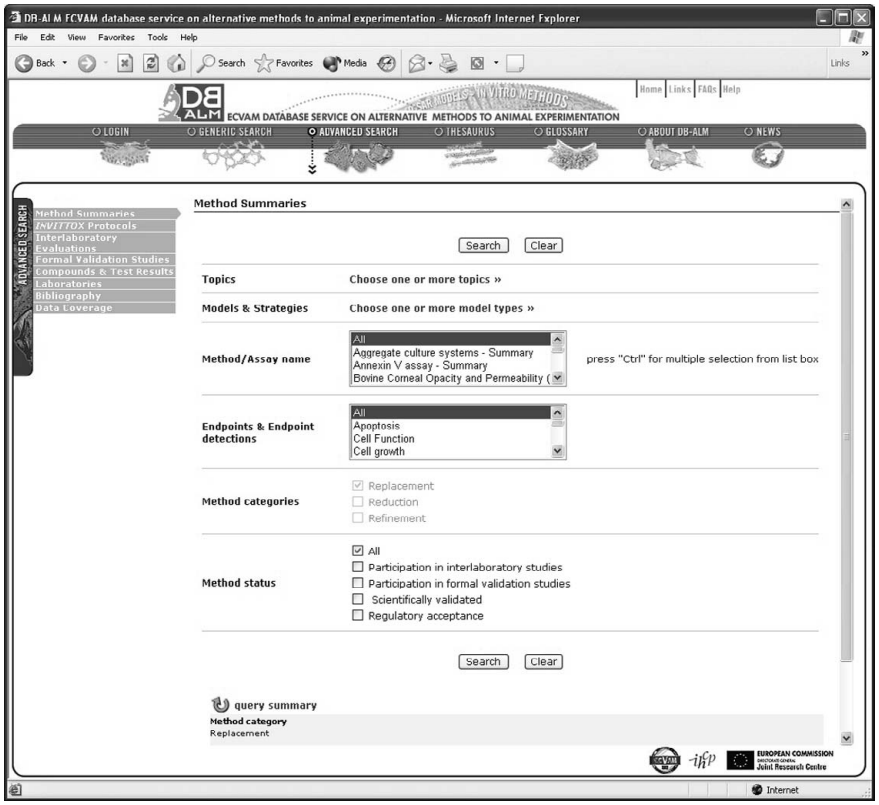


Fig. 3: Method summary search screen

Industry (31%), National Authorities and Commission Services (14%), and others.

Internet addresses

Current DB-ALM version: http://ecvam-sis.jrc.it (until 2006)
ECVAM website: http://ecvam.jrc.it

DB-ALM in 2006

The new version of DB-ALM is already running internally at ECVAM for trial and has been used for the first time also for external data entry by external international experts. In 2006, public access will be opened to this entirely revised Internet version of ECVAM's database service, which will replace the current Internet version known as "SIS (http://ecvam-sis.jrc.it)". A preview of the new homepage, the main portal to all data sectors, together with a few database displays is included hereafter. In addition to the already available *INVITTOX* protocol collection, method-summary descriptions will be included, as well as details on formal validation studies and test results. Furthermore, access to the newly developed data sector on (Q)SARs will also be allowed (fig. 2, 3, 4 and 5).

References

Communication of the European Commission to Council and the European Parliament, Commission of the European Communities, Brussels, 29.10.91 SEC(91)1794 final.

Council Directive 86/609/EEC on the approximation of the laws, regulation and administrative provisions of the Member States regarding the protection of animals used for experimental and other scientific purposes. OJ L Nº 358, 18.12.1986.

Grune, B., Fallon, M., Howard, C. et al. (2004). Report and Recommendations of the workshop "Retrieval approaches for information on alternative methods to animal experiments". *ALTEX 21/3*, 106-127.

Janusch Roi, A. and Balls, M. (1999). The ECVAM Scientific Information Service (SIS). *ALTEX 16/3*, 114-116.

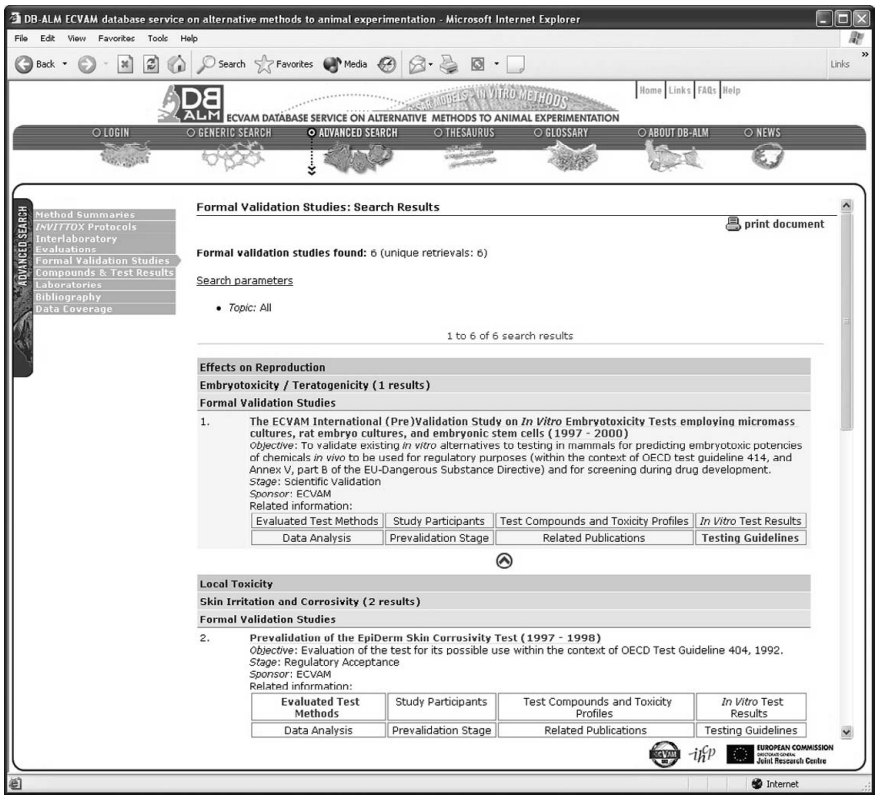


Fig. 4: Method Summary Database: display of search results and access to the related information

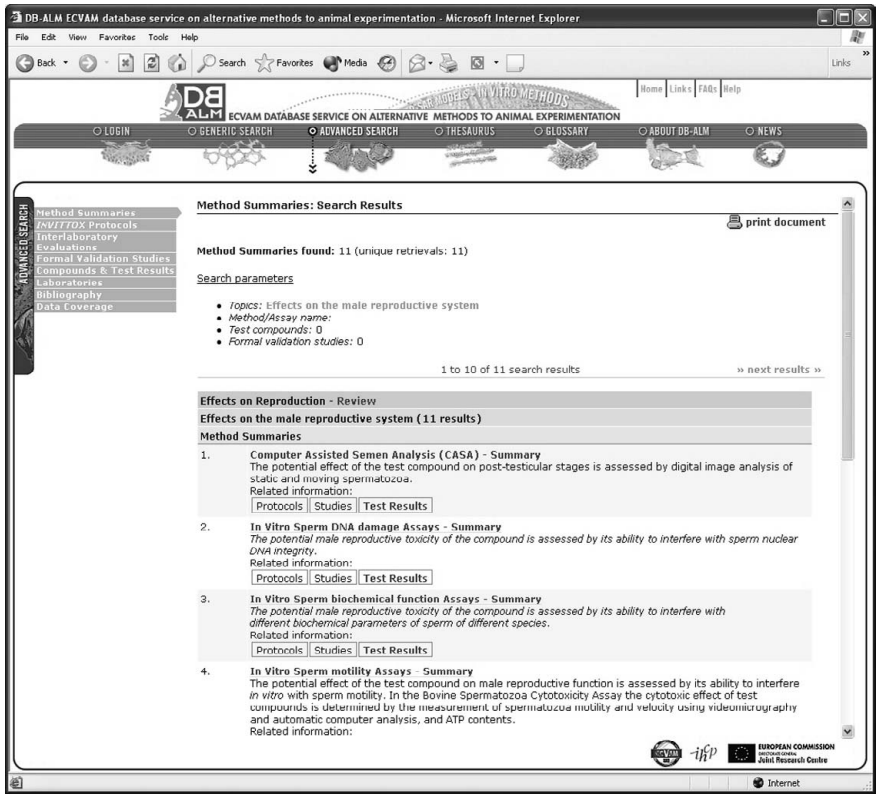


Fig. 5: Validation Study Database: display of search results and access to the related information

Janusch, A., van der Kamp, M. D. O., Bottrill, K. et al. (1997). Current Status and Future Developments of Databases on Alternative Methods. ECVAM Workshop Report 25. *ATLA 25*, 411-422.

Roi, J. A. and Burke, A. (2004). The ECVAM Scientific Information Service. Proceedings of the Fourth World Congress in New Orleans 2003. *ATLA 32, Supplement 1*, 569-572.

Ungar, K. (1993). INVITTOX; an attempt to solve some information problems in in vitro toxicology. *Humane Innovation and Alternatives 7*, 540-543.

Correspondence to
Dr. Annett J. Roi
ECVAM, Institute for Health & Consumer Protection
Joint Research Centre
European Commission
Via E. Fermi 1, TP 580
21020 (VA) Ispra
Italy
e-mail: annett.roi@jrc.it

Session 4.2
Information retrieval – search strategies and search engines

Alternative Search Methods to Retrieve Information on the Web

Hua Florence Chang
National Library of Medicine, National Institutes of Health, U.S. Department of Health and Human Services, Bethesda USA

Summary
It is difficult to retrieve relevant scientific information on alternative methods to animal use. Since 1992, the National Library of Medicine (NLM) has made efforts to improve access to its bibliographic collection on animal use alternatives. In 2001, the NLM began ALTBIB, an online version of previously developed bibliographies. The NLM also added several key journals on alternatives to animal use to MEDLINE/PubMed, and added some relevant new MeSH terms. The ALTBIB search interface has been enhanced to allow searches conducted in MEDLINE/PubMed through a search strategy filter tailored for alternatives. This interface enables users to find the latest citations.

Keywords: alternative methods, database, information retrieval, search strategy, Internet, MEDLINE, PubMed

Introduction

The scientific community is sensitive to concerns about how and why animals are used in biomedical research and testing. Although alternatives to methods based on the use of animals may not satisfy all the requirements of the biomedical research communities, alternatives to the use of vertebrates have been developed and validated. Research on such methodologies is aimed at refining procedures to reduce pain and discomfort, reducing the number of animals required to provide scientifically valuable results, and replacing live vertebrates when an alternative methodology can be verified and validated by the scientific community (the 3Rs).

NLM's ALTBIB (Hudson et al., 2002) is a searchable bibliographic collection on alternatives to animal testing. It includes citations from published articles, books, book chapters, and technical reports from 1992 to 2001. Citations were selected manually after searching various NLM bibliographic databases. The selection of citations involved further review and verification by subject specialists to ensure relevance and quality. The bibliography features citations concerning methods, tests, assays, and procedures that may prove useful in establishing alternatives to the use of intact vertebrates.

In 2001, several enhancements were made to MEDLINE to facilitate the information retrieval of animal alternatives literature. These included the addition of new MeSH main headings on the animal alternatives subjects. Several important journals relevant to animal alternatives were added to the journals indexed for MEDLINE. Since the introduction of these enhancements in MEDLINE and MeSH, ALTBIB has not been updated. Believing that users would be able to search and retrieve directly from MEDLINE/PubMed, it was decided that the effort to manually curate an extensive bibliography was no longer justifiable. Furthermore, the manually created bibliography was limited by the fact that it was never current, but rather became obsolete as soon as produced.

Later study (Grune et al., 2002) showed that the literature relevant to animal alternatives was not easily retrievable. While the new MeSH terms were used, on the whole, to indicate the topics discussed in the citation, many relevant articles would not be

indexed with the new terms. Obtaining relevant information on animal alternatives from MEDLINE/PubMed remains an ongoing problem. Search strategies for animal alternatives have been developed, and incorporated in the ALTBIB search interface. This paper describes the method of developing the search strategy.

Materials and methods

A method was developed to formulate a MEDLINE/PubMed search strategy from MeSH headings assigned in key journal citations. The following steps were taken, in the order indicated:

1. Identify the key journals

MEDLINE/PubMed search was performed using the MeSH heading "Animal Use Alternatives". Search results were downloaded, and the journal title field was extracted, sorted and ranked by frequency. The frequencies were normalised with the total count of citations from each journal.

2. Retrieve all citations from key journals

A MEDLINE/PubMed search with the query using the Journal Title mnemonics "TA" as qualifiers was performed:
Altern Lab Anim[TA] OR ALTEX[TA] OR Toxicol In Vitro[TA] OR ILAR J[TA]

3. Extract and rank MeSH headings by frequency

A Unix shell script was used to extract the MH (MeSH heading) field from all the citations retrieved in steps 1 and 2. The MeSH headings were then sorted and ranked by frequency in descending order.

4. Extract top-ranked relevant MeSH headings

The list of frequency-ranked MeSH headings was examined to find and eliminate terms that were "check tags" (e.g. Human, Animal) or very non-specific terms (e.g. Time Factors, Reproducibility of Results). The top ranking MeSH headings were examined by a subject expert, and descriptors related to animal alternatives were identified.

5. Find co-occurring MeSH headings for general terms

Some of the frequently occurring MeSH headings were quite general. If these MeSH headings were to be included in the search strategy, they needed to be coordinated with other MeSH headings (and qualifiers). To find the coordinating MeSH headings for these general terms, co-occurring MeSH headings were extracted, sorted, and ranked by frequency. For example, the important concept of pain and distress in animals could be expressed with co-ordinated MeSH indexing: (Pain OR Stress, Psychological) AND Animal Welfare.

6. Formulate a search strategy

A search strategy was formed by concatenating the MeSH headings and expressions. In addition to the MeSH headings and expressions, relevant journal titles were also included. All the terms were combined with the Boolean operator OR.

7. Incorporate search strategy to ALTBIB interface

The ALTBIB search interface was then enhanced to allow searches in MEDLINE/PubMed. One or more user-specified query terms are combined with the search strategy with the Boolean operator "AND". Additional search "Limits" were provided to narrow the searches for citations published in 2000-2005, "Animal Use Alternatives" indexed citations, and toxicology subsets. In addition, the "Edit Search Strategy" page allows users to modify or add terms to the search strategy.

Results

The key journals were identified by analysing the search results of "Animal Use Alternatives". The highest frequency journals with citations indexed with "Animal Use Alternatives" included: *Alternatives to Laboratory Animals, ALTEX, Toxicology In Vitro*, and *ILAR Journal* (see tab.1).

Over 5000 citations were retrieved from a search within these 10 key journals.

The citations were downloaded in MEDLINE format, using the MEDLINE/PubMed "Send to File" feature. MeSH headings were extracted from the citations. They were sorted and ranked by frequency of occurrence.

Following the steps outlined above, a collection of high frequency relevant MeSH headings and expressions, and coordinating MeSH headings was developed. To formulate the final search strategy, the MeSH headings and expressions were combined. Key journal titles were also added to ensure the inclusion of citations from these journals. Boolean operator OR bundled the search terms together to form the final search strategy (see tab. 2).

Tab. 1: Frequency journals are indexed with "Animal Use Alternatives"

	Frequency of "Animal Use Alternatives" Index	Total count of citations in this journal	Normalised frequency
Alternatives to Laboratory Animals	234	350	0.668
ATLA Abstracts	8	19	0.421
Society and Animal	4	15	0.267
ALTEX	101	464	0.217
Toxicology *In Vitro*	68	575	0.118
In Vitro & Molecular Toxicology	4	53	0.075
ILAR Journal	23	305	0.075
Developments in Biologicals (Basel)	27	515	0.052
Journal of Applied Animal Welfare Science	3	88	0.034
Developments in Biological Standardisation	99	2983	0.033

Searches in ALTBIB to MEDLINE/PubMed were formulated to combine the search terms with the search strategy. For example, for a search of "Up and Down Procedure", the query to MEDLINE/PubMed was formed as:

"Up and Down Procedure" AND (*In vitro*[mh] OR cell lines[mh] OR cells, cultured OR 3t3 cells OR animal use alternatives OR animal testing alternatives OR animal experimentation OR animal welfare OR laboratory animal science/mt OR animal rights OR animal housing OR animal care committees OR ((pain OR psychological stress) AND animal welfare) OR animal husbandry OR toxicity tests/mt OR toxicology/mt OR tetrazolium salts OR neutral red OR comet assay OR LLNA OR mutagenicity tests OR mutagens/tox OR carcinogens/tox OR teratogens/tox OR animals, laboratory OR QSAR OR structure activity relationship OR ALTEX[TA] OR ILAR J[TA] OR Toxicology[TA] OR Toxicol In Vitro[TA] OR Toxicol Lett[TA] OR Reprod Toxicol[TA] OR Toxicol Sci[TA] OR Food Chem Toxicol[TA] OR Hum Exp Toxicol[TA] OR Regul Toxicol Pharmacol[TA] OR Altern Lab Anim[TA] OR Arch Toxicol[TA] OR Contemp Top Lab Anim Sci[TA] OR Lab Anim NY[TA] OR ATLA Abstr[TA] OR Environ Health Perspect[TA] OR Toxicol Pathol[TA] OR Fundam Appl Toxicol[TA] OR Soc Anim[TA] OR In Vitr Mol Toxicol[TA] OR Dev Biol Basel[TA] OR J Appl Anim Welf Sci[TA] OR Dev Biol Stand[TA] OR J Vet Med Educ[TA] OR (Vaccines AND Animal Use Alternatives) OR (Antibodies, Monoclonal/biosyn AND (Hybridomas OR Bioreactors).

As a result of these efforts, it is now possible to use the ALTBIB search interface to do a selective search for citations in PubMed involving the particular area of interest combined with a broad search strategy to identify those citations of particular interest for alternatives to animal use.

Discussion and conclusion

Efforts by the NLM to improve information about animal alternatives have been going on since 1992. The first efforts resulted in the ALTBIB database, containing over 7,500 citations on tests and methods that refine, reduce, or replace animal experimentation. However, the process used to create this bibliography could not support the increasing amount of information available. Review and verification by subject specialists was necessary to ensure the relevance of the citations. As the explosion of information in the biomedical fields progressed, the human effort involved in the review process became unsupportable.

Furthermore, a major limitation of producing a selective bibliography such as this is that the resulting bibliography is limited by time. It cannot accommodate new materials in a timely manner, but instead remains a static picture, perhaps subject to periodic updates.

In 2001, NLM made several additional efforts to improve access to information about animal alternatives. Knowing that indexing practices could never identify all the articles relevant to animal alternatives, certain concepts seemed of sufficient maturity to add as MeSH headings. These added headings would facilitate indexing of journal citations in which some of the animal alternatives concepts were discussed.

Additionally, after a meeting at NLM, and on the recommendation of representatives from the scientific community, several journals especially focused on animal alternatives were added to those indexed for MEDLINE.

A recent study by ZEBET indicated that the current indexing systems do not provide the required information, since not all of the relevant information is indexed under "alternative methods". The development of suitable search strategies on alternative methods was recommended by representatives of animal welfare information centres.

Search strategy retrieval of animal alternatives information from MEDLINE/PubMed was developed by the NLM ALTBIB team. It has been incorporated into the new ALTBIB interface. The search strategy provides an alternative way of providing relevant information to the animal alternatives community.

In addition to using the search strategy method, there are other methodologies that could be applicable to animal alternatives information retrieval on the Web. Table 3 summarises the various approaches to supporting information access.

Tab. 2: Search strategy terms grouped by categories

Category of concepts	MeSH expressions
In Vitro	(*In Vitro* [MH]) (Cell Lines [MH]) (Cells, Cultured [MH]) (3T3 Cells)
Animal Alternatives	(Animal Use Alternatives) (Animal Testing Alternatives) (Animal Experimentation)
Animal Welfare	(Animal Welfare) (Laboratory Animal Science/methods) (Animal Rights) (Animal Housing) (Animal Care Committees) ((Pain OR Psychological Stress) AND Animal Welfare) (Animal Husbandry)
Methods / Reagents	(Toxicity Tests/methods) (Toxicology/methods) (Tetrazolium Salts) (Neutral Red) (Comet Assay) (LLNA) (Mutagenicity Tests)
Toxicological Concepts	(Mutagens/toxicity) (Carcinogens/toxicity) (Teratogens/toxicity)
Animals	(Animals, Laboratory)
Structure-Activity	(QSAR) (Structure-Activity Relationship)
Journals	(ALTEX) (ILAR J) (Toxicol In Vitro) (Altern Lab Anim)(Contemp Top Lab Anim Sci) (Lab Anim (NY)) (ATLA Abstr) (Environ Health Perspect) (Toxicol Pathol) (Fundam Appl Toxicol) (Toxicology) (Toxicol Lett) (Reprod Toxicol) (Toxicol Sci) (Food Chem Toxicol) (Hum Exp Toxicol) (Regul Toxicol Pharmacol) (Arch Toxicol) (Soc Anim) (In Vitr Mol Toxicol) (Dev Biol Basel) (J Appl Anim Welf Sci) (Dev Biol Stand) (J Vet Med Educ)
Other	(Vaccines AND Animal Use Alternatives) (Antibodies, Monoclonal /biosyn AND (Hybridomas OR Bioreactors)

Tab. 3: Comparison of search methods on the Internet

	Level of Effort	Updates	Scope/Resources	Sustainability
Customised Bibliography	High. Continuously collect and review.	Tends to lag behind.	Single.	Low.
Search Strategy	Low to Moderate. Create search strategy initially, modify as needed.	Retrieves the most . current information	Single.	High.
Information Portal	Moderate. Identify and maintain list of Web resources. Make all participating sites searchable.	Retrieves the most current information.	Multiple Web resources. Only static pages are searchable.	Moderate.
Meta-search	Moderate. Identify and maintain list of Web resources. Configure initially.	Retrieves the most current information.	Multiple Web resources. Both static pages and hidden web (databases) are searchable.	High.

Information Portals

Information portals are collaborations among organisations to develop cross-organisation websites (portals) to provide the breadth of information and services available on a particular topic or audience group. An example of this work is the ALTWEB website (Hakkinen and Green, 2002). ALTWEB, collaborating with member institutions, offers a one-stop shop for animal alternatives information. The information portal provides a starting point, or a gateway, to other related resources. Collaboration among the participating organisations is crucial. While a portal organises information and provides links to the resources of the participating organisations, content or data integration could be a challenge. Information from heterogeneous sources should be made available to applications or users of the portal website.

Meta-Search and Clustering

Meta-search and clustering techniques have been developed using natural language processing methodologies. Target databases or Web information resources can be designated for a custom search. The search results are analysed, and concepts of interest are aggregated and clustered. Users may then refine or modify their queries. For example, at the National Library of Medicine, a meta-search and clustering engine has been developed to perform topic specific searches in areas of AIDS/HIV and toxicology/ environmental health information.

Future

NLM is committed to provide published information on methods that replace, reduce or refine animal use in scientific experimentation. Our goal is to refine search strategies, reduce the number of missed important papers, and replace the legacy information architecture with a new computing paradigm.

References

Grune B., Fallon M., Howard C. et al. (2004). Report and recommendations of the international workshop "Retrieval approaches for information on alternative methods to animal experiments. *ALTEX 21(3)*, 115-27.

Hakkinen PJ., Green DK. (2002). Alternatives to animal testing: information resources via the Internet and World Web Web. *Toxicology, 173(1-2)*, 3-11.

Hudson V., Chang F., Mashayekhi B. (2002). National Library of Medicine's Alternatives to Animal Testing Database on the Web. Congress Proceeding, Fourth World Congress on Alternatives and Animal Use in the Life Sciences.

Correspondence to

Hua Florence Chang
Specialized Information Services Division
National Library of Medicine
National Institutes of Health,
U.S. Department of Health and Human Services
6707 Democracy Blvd, Suite 510
Bethesda, MD 20792 USA
e-mail: changf@mail.nih.gov

Searching Strategies for Detecting Publications on Alternative Methods: A Pilot Study

Barbara Grune, Amrei Schnock, Antje Dörendahl, Susanne Skolik and Horst Spielmann
Federal Institute of Risk Assessment (BfR), Centre for the Documentation and Evaluation of Alternative Methods to Animal Experiments (ZEBET), Berlin, Germany

Summary
As a consequence of the increasing regulatory acceptance of the 3Rs concept of Russel and Burch (1959) on using alternative methods to animal experiments, it is essential that scientists, animal welfare officers, and public policy-makers are able to retrieve relevant, high-quality publications on alternative methods. Efficiently retrieved scientific literature will improve the application of alternative methods. In a pilot study we asked: What are effective search strategies for articles on alternative methods to acute oral toxicity testing? Free text searches (n=28) were performed in databases. The retrieval performance of selected search terms (n=17) was evaluated. We determined the relevance of the articles retrieved and identified a first ranking of search terms and databases for searching for articles on alternative methods used in acute oral toxicity testing. More research will be needed to address other areas of toxicity testing. Recommendations of search strategies to improve the success of searching for articles on alternative methods should include appropriate search terms, phrases and recommendations for databases.

Keywords: animal welfare, animal testing alternatives, databases, searching strategies, search terms

Introduction

In EU member states and in the USA, scientists are obliged by animal welfare legislation not to conduct an animal experiment if another scientifically satisfactory method to obtain the desired information is reasonably and practicably available (Council Directive 86/609/EEC, United States Department of Agriculture, 1997; United States Department of Agriculture, 2000).

To meet the regulatory obligation to use alternatives to experimental animals, scientists should consult literature and other relevant sources on alternatives prior to any experimental study using laboratory animals. It is the responsibility of each scientist to select the most appropriate database to obtain information on alternatives, which have been defined as methods that refine, reduce or replace animal experiments (The 3Rs Concept of Russell and Burch, 1959).

Animal welfare legislation in the European Union and the USA is the basis for information service systems supporting searches for alternative methods. It is essential that scientists, animal welfare officers, and public policy-makers can retrieve relevant, high-quality alternative methods reports. A wide range of information resources, e.g. databases and websites, offer scientific information on alternative methods (Hakkinen et al., 2002). Scientists expect to obtain accurate, complete and relevant information within a short time frame by using online sources, including web-based bibliographic databases and specialised web sites.

Of course, it is in the highest interest of the information provider and information services such as ZEBET, the National German Center for Documentation and Evaluation of Alternative Methods to Animal Experiments, to ensure that the relevant information is readily accessible, especially literature on animal welfare regulations and public health service policies.

However, in our Internet era, the problem is not how to distribute information but rather how to find the right information within an appropriate time frame. This discussion has focused on the problem of information retrieval in the heterogeneous environment of the World Wide Web (Krause, 2002; Adamczak et al., 2002).

In November 2003, at the international workshop, "Retrieval Approaches for Alternative Methods to Animal Experiments" in Berlin, representatives of European and U.S. information centres recommended the development of web-based search tools to improve the information retrieval of alternative methods from the Internet (Grune et al., 2004). Search tools should include search algorithms with suitable search terms. Web-based search tools for alternative methods should allow the use of the most specific search profile reflecting the type of research to which alternative methods are to be applied. Furthermore, searches based on multi-database formats as well as recommendations for databases supporting appropriate searching strategies should be discussed.

Since 1989, ZEBET's information service has responded to inquiries from universities, individual scientists, animal welfare offices and the general public. In 2000, ZEBET introduced AnimAlt-ZEBET, an internet database on alternatives to animal experiments (Grune et al., 2000). Based on ZEBET's experience in documentation, indexing and searching for alternative methods, we endeavoured to answer the question: What are effective strategies to search for articles on alternative methods in databases?

In October 2004, ZEBET began a pilot study focused on information retrieval on alternative methods for testing the acute oral toxicity of chemical and pharmaceutical substances as a

replacement of the "classic" LD_{50}-test. The objective of the pilot study was to translate ZEBET's know-how in documentation, indexing and information retrieval into recommendations for searching strategies to find information on alternative methods in databases.

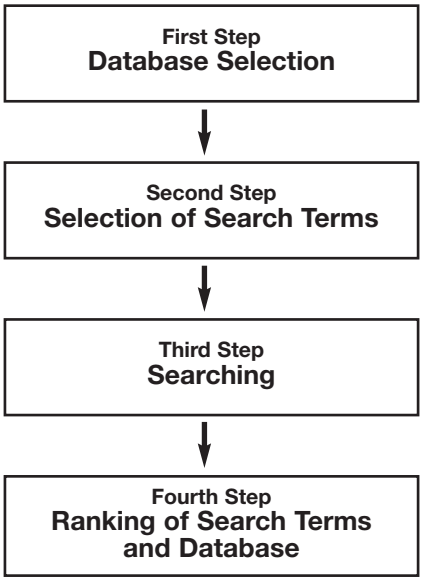
Methods

This study was conducted at the ZEBET, the National German Center for Documentation and Evaluation of Alternative Methods to Animal Experiments, at the German Federal Institute for Risk Assessment between October 2004 and April 2005. A multidisciplinary team of investigators, including scientists and information professionals, undertook this study.

For searching, we used DIMDI, the official host of biomedical databases and provider of medical information systems within the German Federal Ministry of Health, http://www.dimdi.de. DIMDI, the German Institute of Medical Documentation and Information, provides comprehensive scientific data from international publications in the fields of medicine, pharmacology, toxicology and biology. DIMDI offers access to about 70 literature databases, full text databases, as well as factual databases with approximately 100 million documents.

We arranged three database pools based on DIMDI's selection by frequency, two of which consisted of free-of-charge databases and one consisted of commercial, charging databases. Searching was conducted in a superbase mode, which means that all selected databases were searched simultaneously. Altogether, we conducted 28 single free-text searches in preselected databases. Retrieval performance of 17 selected search terms was evaluated. We determined the relevance of the articles retrieved and identified a ranking of search terms and databases for searching for articles on alternative methods to acute oral toxicity testing.

Our investigation consisted of four steps summarised in the following flow chart (see fig. 1):



First Step: Database selection

DIMDI offers a database selection tool called DIMDI Index. Databases were selected based on the frequency of retrieved documents on the comprehensive topic of acute toxicity testing. First, we marked the following relevant subject groups for the index: Human Medicine (all relevant databases), Pharmacology (all relevant databases), Toxicology (all relevant databases), Biotechnology, Genetic Engineering, Medical Devices, Veterinary Medicine.

After marking relevant subject groups, we entered the following search term combination: (acute and tox?)/same sent). This search term combination allowed a broad pre-selection of all databases with information on the topic of acute toxicity. The symbol "?" is a variable truncation. Each question mark is replaced by an arbitrary number of letters. The command ".../same sent" searches for documents in which the search terms "acute" and "tox?" are placed in the same sentence. In that way, we retrieved all documents, in which both search terms were used in close relation. We got an overview of the number of hits resulting from our inquiry in the individual databases, including free-of-charge databases and commercial databases. We selected the following three database pools for searching, based on the retrieved DIMDI Index specific for the topic of acute toxicity.

The first database pool, made up of 11 free-of-charge databases, consisted of a total of 21,014,273 documents on December 9, 2004 and 21,020,203 documents on December 12, 2004. The first database pool contained the following databases: ZT00 – AnimAlt-ZEBET; CL63 – CancerLit; CA66 – CATFILEplus; KL97 – Kluwer-Verlagsdatenbank; ME0A – MEDLINE Alert; ME60 – OLDMEDLINE; ME66 – MEDLINE; SM78 – SOMED; SPPP – Springer-Verlagsdatenbank; TV01 – Thieme Verlagsdatenbank;T165 – XTOXLINE.

The second database pool, made up of 23 free-of-charge databases, consisted of a total of 22,436,740 documents on January 10, 2005. The second database pool contained the following databases: ZT00 – AnimAlt ZEBET; CC00 – CCMed; CL63 – CancerLit; CA66 – CATFILEplus; CDSR 93 – Cochrane Library – CDSR; CDAR 94 – The Cochrane Database of Abstracts of Reviews of Effectiveness; AR 96 – Deutsches Ärzteblatt; GM 03 – gms; GA03 – gms Meetings; HN 69 – HECLINET; INAHTA – NHS-CRD-HTA; KR 03 – Karger Verlagsdatenbank; KL97 – Kluwer-Verlagsdatenbank; LWW04 – LWW-Verlagsdatenbank; MK77 – MEDIKAT; ME0A – MEDLINE Alert; ME60 – OLDMEDLINE; ME66 – MEDLINE; SM78 – SOMED; SPPP – Springer-Verlagsdatenbank; SP97 – Springer-Verlagsdatenbank; TV01 – Thieme-Verlagsdatenbank; T165 – XTOXLINE.

The third database pool, made up of 21 commercial, charging databases, consisted of a total of 69,741,166 documents on February 7, 2005. The third database pool contained the following databases: CB85 – AMED; AN83 – Adis Newsletters; BA70 – BIOSIS Previews; CV72 – CAB Abstracts; CCTR93 – Cochrane Library – Central; BD82 – Derwent Biotechnology Resource; DH64 – Derwent Drug Backfile; DD83 – Derwent Drug File; EA08 – EMBASE Alert; EM74 – EMBASE; EB94 – Elsevier BIOBASE; FS69 – FSTA; AZ72 – GLOBAL Health; HA85 – HAD; HT83 – IHTA; IA70 – IPA; II78 – ISTPB + ISTP/ISSHP; MT68 MEDITEC; NHSEED – NHS-EED; IS74 – SciSearch; TB69 – TOXBIO.

Second Step: Selection of search terms

We searched a total of 17 terms: one specific search term for the topic of our investigation concerning acute oral toxicity, 10 general terms for alternative methods and animal welfare, and 6 relevant biomedical search terms. We selected exclusive search terms in English, since the majority of relevant literature is published in English, including abstracts and keywords.

The selection of search terms was based on the terms of the Medical Subject Headings thesaurus (MeSH), the index terms of AnimAlt-ZEBET-Database, and definitions of professional associations.

AnimAlt-ZEBET documents have been evaluated by ZEBET's staff according to the 3Rs principle; the documents provide an assessment of the current stage of development, validation and acceptance of the methods for regulatory purposes. Each document is characterised by specific keywords and includes abstract and bibliographic references. ZEBET's index terms correspond to those of MeSH. MeSH is a controlled vocabulary produced by the National Library of Medicine (NLM) and used for indexing, cataloguing, and searching for biomedical and health-related information and documents.

● Search term for the topic of acute oral toxicity

ZEBET's pilot study focused on information retrieval on alternative methods for testing the acute oral toxicity of chemical and pharmaceutical substances as a replacement of the classic LD_{50}-Test. This topic was chosen because there are clear definitions for acute toxicity as well as for acute oral toxicity. In addition, there are well-published alternative methods and five documents in the AnimAlt-ZEBET database.

The MeSH term "Acute Toxicity Tests" is defined by the NLM (2005) as follows: Experiments designed to determine the potential toxic effects of one-time, short-term exposure to a chemical or chemicals.

Acute toxicity studies include different test procedures, depending on the type of administration, e.g. acute oral toxicity, acute dermal toxicity or acute inhalation toxicity. As mentioned above, we searched literature for the topic of alternative methods for acute oral toxicity testing. The Organisation for Economic Co-operation and Development (OECD) defines "Acute oral toxicity" as: the adverse effects occurring within a short time of oral administration of a single dose of a substance or multiple doses given within 24 hours (OECD, 1993). The principle of the LD_{50} test is to dose groups of animals with a single dose of a test substance at concentrations expected to cause death in at least a fraction of the animals dosed (OECD, 2002).

AnimAlt-ZEBET contains five alternative method documents for acute oral toxicity testing, in which the term "acute oral toxicity" is used (AnimAlt-ZEBET, 2005):

- Determination of the approximative LD_{50} for testing the acute toxicity of pharmaceutical substances as a replacement of the classic LD_{50}-test
- Up-and-down procedure for testing the acute oral toxicity of chemicals with a significantly reduced number of solely female rats as a replacement of the classical LD_{50} test
- Fixed Dose Procedure for testing the acute oral toxicity of chemical substances as a replacement of the classical LD_{50} test
- Acute Toxic Class Method for testing the acute oral toxicity of chemicals as a replacement of the classical LD_{50} test
- Determination of the starting dose for acute toxicity (LD_{50}) testing of chemical substances by applying linear regression models of "The Registry of Cytotoxicity"

● Search terms for alternative methods and animal welfare

The selection of search terms for alternative methods and animal welfare is based on ZEBET's investigation of indexing systems of alternative methods in established literature databases, e.g. MEDLINE, Embase, Agris, CAB Abstracts, and Agricola (Grune et al., 2004; Meißner, 2002). We identified the following search terms for alternative methods and animal welfare corresponding to the 3Rs Concept of Russel and Burch:

- alternatives to animal testing
- animal testing alternative, animal testing alternatives
- animal testing reduction
- animal testing refinement
- animal testing replacement
- animal use alternatives
- animal use reduction
- animal use replacement
- animal use refinement
- animal welfare

The MeSH-Term "Animal Testing Alternatives" (1985-2000) is defined as: Procedures, such as tissue culture, mathematical models, etc., when used or advocated for use in place of the use of animals in research or diagnostic laboratories (National Library of Medicine, 2005). Since 2000, the term "Animal Use Alternatives" is used by the NLM as: Alternatives to the use of animals in research, testing, and education. The alternatives may include reduction in the number of animals used, replacement of animals with a non-animal model or with animals of a species lower phylogenetically, or refinement of methods to minimise pain and distress of animals used.

The MeSH explains the term "Animal Welfare" as: The protection of animals in laboratories or other specific environments and the promotion of their health through better nutrition, housing, and care. This may be carried out through legislation or regulation.

● Relevant biomedical search terms

In addition, we selected the following search terms, which are used in the scientific literature and in AnimAlt-ZEBET documents to describe alternative methods for acute toxicity testing:

- cell culture, cell cultures
- *in vitro*
- model
- method, methods
- procedure, procedures
- approach , approaches

Third Step: Searching

● Search modus

We used DIMDI ClassicSearch, the expert search modus for command language professionals. DIMDI ClassicSearch offers options to formulate search queries and provides comprehensive

access to all of DIMDI's databases. For example, we used truncated and qualifying search terms, combined search terms with logical operators, and modified their sequence using parentheses. Truncated search terms allowed to search several terms with one command, e.g. "animal test?", which covers animal testing alternative, animal testing alternatives, animal testing reduction, animal testing refinement, animal testing replacement.

• Search mode

We searched in free text search mode to find all available documents containing information on alternative methods in any accessible text field of the documents, e.g. title, keywords, abstract.

Fourth Step: Ranking of search terms and databases

The retrieval performances of selected search queries were evaluated according to the relevance of the content of documents retrieved for the topic of alternative methods for acute oral toxicity testing. Relevance has been evaluated and was confirmed by ZEBET's staff in comparison to AnimAlt-ZEBET documents, including references for alternatives for acute oral toxicity testing. From the number of retrieved relevant documents, we identified a first ranking of search terms and databases to search for articles on alternative methods to acute oral toxicity testing.

Tab. 1: Searching in the first database pool – 11 databases free-of-charge – December 2004

Search number	Date	Number of documents in total	Search queries acute oral tox AND ...	Retrieved documents	Relevant documents
1	09.12.2004	21,014,273	Acute Oral Tox? AND ALTERNATIV?	18	17
2	09.12.2004	21,014,273	Acute Oral Tox? AND ANIMAL TEST?	15	9
3	09.12.2004	21,014,273	Acute Oral Tox? AND ANIMAL USE	2	2
4	09.12.2004	21,014,273	Acute Oral Tox? AND ANIMAL WELFARE	10	10
5	09.12.2004	21,017,709	Acute Oral Tox? AND (CELL# AND CULT?/SAME SENT)	35	1
6	09.12.2004	21,017,709	Acute Oral Tox? AND VITRO	39	3
7	09.12.2004	21,017,709	Acute Oral Tox? AND MODEL?	31	10
8	12.12.2004	21,020,203	Acute Oral Tox? AND METHOD?	30	8
9	10.12.2004	21,020,203	Acute Oral Tox? AND PROCEDURE?	30	17
10	12.12.2004	21,020,203	Acute Oral Tox? AND APPROACH?	16	6
				226 Documents	**83 Documents**

Tab. 2: Searching in the second database pool – 23 databases free-of-charge – January 2005

Search number	Date	Number of documents in total	Search queries ((acute AND tox?)/ same sent) AND oral/same sent AND ...	Retrieved documents	Relevant documents
1	10.01.2005	22,436,740	((acute AND tox?)/same sent) AND oral/same sent AND alternativ?	29	25
2	10.01.2005	22,436,740	((acute AND tox?)/same sent) AND oral/same sent AND animal test?	27	11
3	10.01.2005	22,436,740	((acute AND tox?)/same sent) AND oral/same sent AND animal use	2	2
4	10.01.2005	22,436,740	((acute AND tox?)/same sent) AND oral/same sent AND animal welfare	13	13
5	10.01.2005	22,436,740	((acute AND tox?)/same sent) AND oral/same sent AND vitro	15	3
6	10.01.2005	22,436,740	((acute AND tox?)/same sent) AND oral/same sent AND model?	13	5
7	10.01.2005	22,436,740	((acute AND tox?)/same sent) AND oral/same sent AND method?	29	13
8	10.01.2005	22,444,599	((acute AND tox?)/same sent) AND oral/same sent AND procedur?	18	17
9	10.01.2005	22,436,740	((acute AND tox?)/same sent) AND oral/same sent AND approach?	25	5
				171 Documents	**94 Documents**

Tab. 3: Searching in the third database pool – 21 commercially charging databases – February 2005

Search number	Date	Number of documents in total	Search queries ((acute and tox?)/same sent) AND oral/same sent AND …	Retrieved documents	Relevant documents
1	07.02.2005	69,741,666	((acute AND tox?)/same sent) AND oral/same sent AND alternativ?	64	34
2	07.02.2005	69,741,666	((acute AND tox?)/same sent) AND oral/same sent AND animal test?	21	9
3	07.02.2005	69,741,666	((acute AND tox?)/same sent) AND oral/same sent AND animal use	8	8
4	07.02.2005	69,741,666	((acute AND tox?)/same sent) AND oral/same sent AND animal welfare	15	13
5	07.02.2005	69,741,666	((acute AND tox?)/same sent) AND oral/same sent AND vitro	13	6
6	07.02.2005	69,741,666	((acute AND tox?)/same sent) AND oral/same sent AND model?	23	8
7	07.02.2005	69,741,666	((acute AND tox?)/same sent) AND oral/same sent AND method?	72	18
8	07.02.2005	69,741,666	((acute AND tox?)/same sent) AND oral/same sent AND procedur?	14	12
9	07.02.2005	69,741,666	((acute AND tox?)/same sent) AND oral/same sent AND approach?	23	9
				253 Documents	**117 Documents**

Results

Number of retrieved and relevant documents

Results are presented on the number of retrieved and relevant documents for each search query and are summarised in three tables. The tables contain the search queries, number of the retrieved and relevant documents. The search profile was recorded for each search query.

Searching in the first database pool started with the fixed, and consequently more restricted, search query "acute oral tox?" (see tab. 1). Searching in the second and third database pool started with search term combination "((acute and tox? and oral)/same sent)", which allowed wider searching (see tab. 2 and 3). Next, search terms for acute oral toxicity were combined with terms for alternative methods and animal welfare.

Searching in the first database pool retrieved 226 documents, 83 of these were classified as relevant. In comparison, searching in the second database pool retrieved 171 documents of which 94 documents were relevant. Both database pools contained free-of-charge databases with a total number of documents between ca. 21 and 22 million. In general, the number of retrieved documents is very low in relation to the total number.

The third database pool contained exclusively databases charging a fee. In this pool, the number of ca. 70 million documents is clearly higher than in the pools of databases that offer

Tab. 4: Ranking of search terms used for searching in the first database pool – 11 databases free-of-charge

Ranking of search terms	Number of relevant documents
1. alternativ?	17
procedure?	17
2. animal welfare	10
model	10
3. animal test?	9
4. method?	8
5. approach?	6
6. *vitro*	3
7. animal use	2
8. cell#	1
cult?	1

Tab. 5: Ranking of search terms used for searching in the second database pool – 23 databases free-of-charge

Ranking of search terms	Number of relevant documents
1. alternativ?	25
2. procedur?	17
3. animal welfare	13
method?	13
4. animal test?	11
5. approach?	5
model?	5
6. *vitro*	3
7. animal use	2

Tab. 6: Ranking of search terms used for searching in the third database pool – 21 commercially charging databases

Ranking of search terms	Number of relevant documents
1. alternativ?	34
2. method?	18
3. animal welfare	13
4. procedur?	12
5. approach?	9
animal test?	9
6. animal use	8
model?	8
7. *vitro*	6

Tab. 7: Ranking of databases in the first database pool – 11 databases free-of-charge

Ranking of databases	Number of relevant documents
1. MEDLINE, ME66	43
2. AnimalAlt-ZEBET, ZT00	29
3. SOMED, SM78	9
4. XTOXLINE, TI65	2

Tab. 8: Ranking of databases in the second database pool – 23 databases free-of-charge

Ranking of databases	Number of relevant documents
1. MEDLINE, ME66	52
2. AnimalAlt-ZEBET, ZT00	25
3. MEDLINE Alert, ME0A	7
4. XTOXLINE, TI65	5
5. CancerLit, CL63	3
6. SOMED, SM78	2

Tab. 9: Ranking of databases in the third database pool – 21 commercially charging databases

Ranking of databases	Number of relevant documents
1. BIOSIS Previews, BA70	49
2. EMBASE, EM74	26
3. CAB Abstracts, CV72	14
4. ISTPB + ISTP/ISSHP, II78	11
5. SciSearch, IS74	11
6. MEDITEC, MT68	4
7. Elsevier BIOBASE, EB94	2

free access. Searching in this pool retrieved 253 documents, 117 documents were evaluated as relevant. Therefore, the number of retrieved relevant documents in commercial, charging databases was higher than in free–of-charge databases.

Ranking of search terms

We have drawn up a first ranking of search terms and databases for searching for articles on alternative methods for acute oral toxicity testing based on the total number of relevant documents retrieved (see tab. 4-6).

Tables 4 and 5 show that the search terms “alternativ?” and “procedure?” provided the most relevant documents in free-of-charge databases. Search terms “alternativ?” and “method?” retrieved the most relevant documents in commercial, charging databases. Search terms “alternativ?”, “procedure?”, and “method?” retrieved more relevant documents than the term “vitro” in all database pools.

Ranking of databases

The number of total relevant documents was determined in relation to the search query as well as in relation to each database used in our study. These numbers provide a ranking of the databases (see tab. 7-9).

Six of 23 free-access databases can be recommended for searching for alternative methods for acute oral toxicity testing. These are MEDLINE, MEDLINE Alert, AnimAlt-ZEBET, XTOXLINE, Cancerlit and SOMED.

In addition, 7 of 21 commercial, charging databases can be recommended, i.e. BIOSIS Previews, EMBASE, CAB Abstracts, ISTPB + ISTP/ISSHP, SciSearch, MEDITEC, and Elsevier BIOBASE. Searching in these databases retrieved quantitatively more relevant documents than searching the databases that do not charge for use.

Discussion and conclusion

Our study has evaluated a step-by-step procedure for searching for documents on the topic of acute oral toxicity testing. We carefully selected databases and search terms based on our own AnimAlt-ZEBET database and the internationally accepted MeSH Thesaurus. The relevance of the retrieved documents was evaluated by scientists in comparison with documents and literature from AnimAlt-ZEBET. The results were summarised in a ranked order of search terms and databases.

It became evident that it is useful to start searching with broad search term phrases. Ranking of terms for the next steps of searching differs between databases that do charge for access or do not charge. For example, searching in free-of-charge databases using the terms “alternativ?” and “procedure?” can be recommended. In comparison, the terms “alternativ?” and “method?” can be recommended for commercial, charging databases.

Search terms “alternativ?”, “procedure?” as well as “method?” retrieved more relevant documents on alternative methods to acute oral toxicity testing than the term “*vitro*” in all databases. This result confirms that most alternative methods for testing acute oral toxicity still use animals. Although alternative methods may require the use of animals, the numbers of animals needed for any of the alternatives are drastically reduced. Moreover, one of the alternative tests does not require the death of animals as an endpoint (OECD, 2002).

The database ranking confirmed that MEDLINE is the first choice for searching on the topic of acute oral toxicity testing in databases that do not charge. However, the other ranked databases should be included in searching, because these databases cover a different spectrum of literature and therefore compliment each other. Furthermore, databases that charge for access, such as BIOSIS and EMASE, significantly improve information searching. Each of these databases has its own profile and contains relevant references not available in the other.

Our pilot study focused on one topic, acute oral toxicity testing, and was limited to general terms describing alternatives. We did not search for a specific method, such as the acute toxic class method. We searched very broadly and therefore remained only on the surface of this topic. However, as always, our recommendations offer a well-proven lead-in to searching.

Further investigations are necessary and should be extended to

other topics, e.g. production of monoclonal antibodies or food hygiene. Also, studies in the Internet with search engines such as Google.Scholar or Scirus (Elsevier) would be important. In addition, collaboration is needed among information specialists for alternative methods to develop standard terminology and standard protocols for searching strategy recommendations.

References

Adamczak, W., Nase, A. (2002). *Gaining Insight from Research Information*. 6th International Conference on Current Research Information Systems. Kassel: University Press, 2002, p. 9.

AnimAlt-ZEBET (2005). *ZEBET-Datenbank über Alternativen zu Tierversuchen im Internet (AnimAlt-ZEBET)*. Available from: http://www.bfr.bund.de/cd/1508.

Council Directive 86/609/EEC of 24 November 1986 on the approximation of laws, regulations and administrative provisions of the member States regarding the protection of animal used for experimental and other scientific purposes. *Official Journal L 358, 18/12/1986* P. 0001-0028.

Grune, B., Fallon, M., Howard, C. et al. (2004). The Report and Recommendations of the International Workshop "Retrieval Approaches for Information on Alternative Methods to Animal Experiments" held in Berlin in 2003 at ZEBET, Centre for Documentation and Evaluation of Alternative Methods to Animal Experiments, in the BfR, the German Federal Institute for Risk Assessment. *ALTEX 21, 3*, 115-127.

Grune, B., Herrmann, S., Dörendahl, A. et al. (2000). Die ZEBET-Datenbank über Alternativen zu Tierversuchen im Internet – ein konkreter Beitrag zum Tierschutz von Versuchstieren. *ALTEX 17, 3*, 127-133.

Hakkinen, P. J. and Green, D. K. (2002). Alternatives to animal testing: information resources via the internet and world wide web. *Toxicology 173*, 3-11.

Krause, J. (2003). *Standardization, heterogenety and the quality of content analysis: a key conflict of digital libraries and its solutions*. World Library and Information Congress: 69th IFLA General Conference and Council, 1-9 August 2003, Berlin. Available from: http://www.ifla.org/IV/ifla69/papers/085g-Krause.pdf.

Meißner, Martina (2002). *Vergleichende Untersuchung von Thesauri, Schlagwortlisten und Klassifikationen im Hinblick auf ihre Anwendung für die Indexierung von Informationen zur Thematik "Alternativmethoden zu Tierversuchen" und auf angrenzenden Gebieten*. Diplomarbeit, Fachbereich Informationswissenschaften, Fachhochschule Potsdam, 2002.

National Library of Medicine (2005). *Medical Subject Headings*. Bethesda USA: U.S. National Library of Medicine. Available from: http://www.nlm.nih.gov/mesh/meshhome.html.

Organisation for Economic Co-operation and Development (2002). *OECD Test Guideline 401 will be deleted: A Major Step in Animal Welfare: OECD Reaches Agreement on the Abolishment of the LD_{50} Acute Toxicity Test*. Available from: http://www.oecd.org/document/52/0,2340,en_2649_201185_2752116_1_1_1_1,00.html.

Organisation for Economic Cooperation and Development (1993). *OECD Guideline 401 "Acute Oral Toxicity"*, Adopted: 24 Februrary 1987. OECD Guidelines for the Testing of Chemicals, Section 4: Health Effects. Fifth Addendum to the OECD Guidelines for the Testing of Chemicals, 1993, Paris.

Russell, W. M. S., Burch, R. L. (1959). *The principles of humane experimental technique*. 238 pp. London: Methuen & Co. Ltd.

United States Department of Agriculture, Animal and Plant Health Inspection Service, Animal Care Policy Manual, *Policy #11 – Painful/Distressful Procedures – April 14, 1997*. Available from: http://www.aphis.usda.gov/ac/policy/policy11.pdf.

United States Department of Agriculture, Animal and Plant Health Inspection Service, Animal Care Policy Manual, *Policy #12 – Written Narrative for Alternatives to Painful Procedures – June 21, 2000*. Available from: http://www.aphis.usda.gov/ac/policy/policy11.pdf.

Acknowledgment

The pilot study was funded by the German Federal Institute for Risk Assessment. The authors are grateful to Professor Dr. Dr. Andreas Hensel, President of the BfR, and Mrs. Christel Zimmermann, Scientific Coordinator of the BfR, for their support.

Correspondence to

Dr. Barbara Grune
Federal Institute of Risk Assessment (BfR)
Centre for the Documentation and Evaluation of Alternative Methods to Animal Experiments (ZEBET)
Diedersdorfer Weg 1
12277 Berlin
Germany
e-mail: grune.zebet@bfr.bund.de

The Animal Welfare Information Center: Helping the Community Meet the Information Requirement of the Three Rs

Jean A. Larson
Animal Welfare Information Center, National Agricultural Library, Baltimore, USA

Summary
The following paper presents a brief overview of the legal requirement of the U.S. Animal Welfare Act regarding the search for alternatives. The establishment of the Animal Welfare Information Center at USDA's National Agricultural Library is described. The AWIC approach to searching bibliographic databases for the 3Rs is outlined and examples of the approach, terminology, database choices, and the development of the search strategy are defined and illustrated. A sample search for alternatives in training protocols for trauma life support is provided and discussed.

Keywords: USDA's Animal Welfare Information Center, information services, US Animal Welfare Act, searching the literature for the 3Rs

Introduction

More that 15 years ago, the Animal Welfare Information Center (AWIC) of USDA's National Agricultural Library (NAL) provided time, enthusiasm and monetary support for "The World Congress on Alternatives and Animal Use in the Life Sciences: Education, Research, Testing". The Congress took place on November 14-19, 1993 in Baltimore, Maryland, USA. Obviously, there was and still continues to be enthusiasm for international dialogue on the important topic of how animals are used in education, research and testing.

For 19 years, AWIC staff has worked hard to meet the Congressional mandate to collect and produce a variety of information products, refine searching approaches, and to share knowledge regarding animal care, environmental enrichment and the "alternative search". Today's presentation will cover three basic topics: 1) a very brief overview of the U.S. Animal Welfare Act (U.S. Government Printing Office, 1985) requirements as related to alternatives; 2) A brief overview of the Animal Welfare Information Center; and 3) the AWIC approach to searching literature and other resources for alternatives.

The US Animal Welfare Act (AWA)

The first Laboratory Animal Welfare Act, U.S. Public Law 89-544, was enacted in 1966 with the purpose of preventing pet dogs from being stolen and ending up being used in biomedical research. The law granted the regulatory authority to the Department of Agriculture (The Animal and Plant Health Inspection Service, Animal Care (APHIS, AC)) where a programme was developed to write the regulations to interpret the Act and set up a nationwide inspection system.

In the third amendment in 1985, sweeping changes were made in the law. It provided more protection to animals in research, to exhibition animals and pets, considered humane transport and added more protection of owners from pet theft (U.S. Government Printing Office, 1985) In this presentation the three new mandates relevant to this presentation are: 1) the role of the Institutional Animal Care and Use Committee (IACUC) and the review of research protocols for four animal use concerns; 2) the need for reviews of the scientific literature to prevent unintended duplication, reduce animal pain and distress, and to use improved methodologies; 3) the establishment of an information service at the National Agricultural Library (NAL) to help the regulated community meet information requirements/needs. The service is the Animal Welfare Information Center programme.

Animals covered. The US AWA includes all warm blooded vertebrates with the following exclusions: 1) *Rattus* and *Mus* species; 2) birds bred for research; 3) farm animals used in agricultural research and teaching; 4) horses not used in biomedical research (i.e. rodeos, racing, pony rides, etc.); and 5) retail pet trade except for exotic species and pocket pets.

Activities regulated. Activities regulated under the AWA include the following: animal breeders and dealers, animals used in biomedical research and testing, higher education, in transport situations, and in exhibits (zoos, circuses, and marine mammals).

Alternatives search mandates. The AWA mandates principal investigators (PIs) to consider and document their process for the consideration of "alternatives" "to any procedure that would reasonably be expected to cause more than slight or momentary pain or distress in a human being". PIs are also required to write an assurance statement that an alternatives search or a consulta-

tion with a recognised expert was done for the protocol. The written assurance statement is to include the following: 1) consideration of alternatives based on the 3Rs of Russell and Burch (1959); 2) sources consulted; and 3) databases searched with names, search date and time period covered, keywords and/or search strategy used. The requirements are detailed in Policy # 12 of the APHIS, AC policy manual. (http://www.aphis.usda.gov/ac/policy/policy12pdf.)

The IACUC is responsible for reviewing all institutional protocols that use live animals to ensure the four following concerns are addressed: 1) painful or distressful procedures performed with analgesics or anaesthetics; 2) the research animal is used for only one major operative procedure unless justified; 3) ensure the PI has explored alternatives to painful or distressing procedures; 4) review the PI's written assurance that the research is not unnecessarily duplicative.

The Animal Welfare Information Center (AWIC)

Congressional mandates for AWIC. The US Congress directs the AWIC programme to provide information on the following topics: 1) information that can be used to train employees to use animals appropriately and treat them humanely; 2) aid in providing information that can prevent unintended duplication of research; 3) information that supports the various intents and requirements of the Animal Welfare Act; and 4) improved methodologies of experimentation that reduce animal use, minimise pain and distress, and are less invasive and distressful to the animals.

Services and information products offered by AWIC. In the 19 years since the Center was established, it has provided a variety of services and information products: reference services that include the alternatives search, training via workshops, presentations and exhibiting at conferences, authoring articles, and producing many types of topical publications within the Congressional mandate.

AWIC's user community. Those regulated under the AWA make up 50% of the user community. Other users include local, state and federal government officials; educators; students; libraries; organisations; the media; and the general public. AWIC staff also provide information to researchers around the globe. The following statistics give the reader some idea of the Center's usage. In the last 10 months, the Center responded to ~2,000 reference requests, distributed >25,000 information products, exhibited at 10 meetings, and conducted 15 formal training events.

The AWIC website. Most information products/materials produced during the 19 years have been converted to HTML documents and are available on the AWIC Website <http://www.nal.usda.gov/awic/>. It is a large Website with thousands of pages of content and carefully selected links. Briefly, the Website contains resources that address lab animals, farm animals, animals in exhibits and wildlife, and companion animals. Topics covered include laws and regulations, various techniques and methods, animal diseases, animal models of disease, animal care, addressing pain and distress, environmental enrichment, species-specific information resources, database links, and the "alternative search". From October 2004 to July 2005, there were 2,239,485 hits and 261,729, 248 kBytes were sent. Many users are from outside the US.

Searching for alternatives: The AWIC approach

Reasons for searching for alternatives

In addition to the legal requirements, there are certainly many persuasive reasons for trying to find alternatives to the use of animals in painful procedures. 1) For example, to gain AALAC International accreditation, an institution needs to follow the laws of the "Guide for the Care and Use of Animals". 2) There are societal pressures to make research and testing more pain and distress free by moving to lower species of animals or using no animals at all. 3) There are humane reasons based on ethical concerns about how animals are used and whether they should be granted "rights". 4) There are economic reasons, as animal-based research is expensive, i.e. the costs of: the animals, the facility requirements, the equipment including caging, cleaning and maintenance, feed, and specialised personnel. 5) Science-based reasons are also compelling and are addressed separately below.

At least nine scientific reasons that come rather quickly to mind. 1) Importantly, to perform up-to-date research, a PI needs to stay current in his/her research area. Obviously then, the search should be done *before* a new protocol is written. 2) A good search will address the duplicative research question. 3) Selection of a broad spectrum of databases will allow for a check of unusual sources. 4) The likelihood of using the appropriate numbers of animals increases. 5) We have seen people find new insights leading to new approaches for their research, possibly revealing methods of using other animals and non-animal models. 6) If caretakers and veterinarians understand the physiology, proper handling, training, enrichment possibilities, instinctive behaviour, etc., the animals will be better research subjects. 7) Researching the compounds in a protocol will reveal whether the compounds are appropriate or whether there are confounding factors in the protocol. 8) There is the potential to use less painful and invasive procedures, which leads to better data. 9) The search may reveal that collaborations at institutional, national and international levels may lead to better research and more efficient use of animal, human, facility and monetary resources.

Building blocks for a search

Two important building blocks used to interpret the concepts and realities of the protocol to construct an effective literature search are 1) terminology and 2) information resources.

Terminology – general

Considerations of terminology should include but not be limited to protocol concepts, the general area of study, compounds, techniques, equipment, scientific terms, terms relevant to the 3Rs, Boolean logic operators, limit terms, document types, etc. The fol-

lowing examples will hopefully be useful in understanding the terminology-based decisions needed. Note that in the examples below, "like terms" are put together and the "?" indicates truncation of the word to allow variations in the word endings. (The author recognises the "?" is not used universally in database platforms.)

- Examples: hypothesis related (cardiac surgery); generic and trade names (xylazine, rompun); acronyms (BSE, cns, mab); synonyms (kuru, prion?, scrapie,); truncation examples (pig? stud?). Spelling is a critical concern as databases contain information in various spellings: American (estrous, estrus, labor) and English (oestrous, colour, labour). Variations in words need to be thought about (tse, transmissible spongiform encephalopathies). Those annoying abrv.'s (abbreviations) can make life difficult. Many taxonomic names are in Latin and if there has been revision in the taxonomic name of an organism, the searcher will need to use the new and the old names.

Terminology – alternatives related

The terminology for this part of the search consists of examples of terms that address reduction and refinement as well as a set of terms that address replacement.

- Examples of reduction and refinement might include the following: analges? or painkiller or anodyne; anesthe? or anasthe? or anaesthe?; technique? or method(s) or procedur?; monitor? device? or evaluat?; restrain? or immobili? or restraint? or restrict?; advers?; positive reinforce? or animal training. Note: Most terms used here are developed from the area of study in the protocol!
- Examples of replacement terminology might include the following: model? or artificial, or *vitro*, or culture?; insect? or invertebrate? or fish, or cephalopod?; simulat?, or digital, or interact? or virtual or mannequin? or manikin or model, or assay(s); tissue? or organ? Note: there are numerous places to find possible alternative terms, but the Indexing Branch of the National Agricultural Library that produces the database AGRICOLA has also developed a mini-thesaurus of alternative-oriented terminology. It is called the Animal Use Alternatives Thesaurus. The mini-thesaurus can be used both for retrieval as well as for assigning keywords to an article. Look for this specialised thesaurus under the "Literature and Databases" section of the AWIC website.

Terminology – Boolean and other

There are 3 Boolean logic terms used for sophisticated database retrieval. They are OR, AND and NOT. "OR" tells the system to try and retrieve citations that have at least one of the words in the search statement. "AND" directs the database system to select citations that contain both terms connected by the "AND". "NOT" will eliminate a term or a group of terms from the possible retrieved items.

Other terms can be used to define from which fields the terms should be retrieved. For example, you can limit the retrieval to only the title field (ti), the descriptors (de), or the identifier (id). In the example note application of the topics discussed above. The "S" below stands for select in the DIALOG system.

? Truncation
  --S behav? = behave, behaves, behaviour, behavior, etc.
OR Select all citations with at least ONE word in the set
  --S swine or pig or pigs or porcine
AND Select all citations containing both words
  --S swine and euthan?
NOT Eliminates a term or group
  --S (pig or swine or porcine) not guinea
TI, DE, ID Limit term/s to title, descriptors, or identifiers.
  --S (dog or dogs or cani?)/ti, de, id

Searchable Information resources

The decision on where to retrieve information is critical! Many people do not know how much content varies between databases! When choosing a database resource, be aware of the subject coverage, years of coverage, selection criteria for inclusion, types of materials indexed (i.e. journal, books, reports, conference proceeding, patents, editorial comments, etc.) to build the resource. Such information is usually quite easily found, but requires some research. For example, MEDLINE and the European biomedical database EMBASE are very biomedically oriented. Citations are selected from 4-5000 journals. AGRICOLA, CAB, BIOSIS, and many other databases select from a broader base of resources – books, conference proceedings, reports, electronic resources, government reports, etc. and have a broader subject coverage. Note that such information is also useful for authors wanting a better idea of which database(s) index which resources.

Formulating a database search

The following illustrates the application of all the information given above as building blocks to answer a request often received at AWIC. Note: Refer to Jensen (2005) "Worksheet and Instructions for Alternatives Literature Searching." (This search was developed for attendees to the AWIC workshop.) The worksheet can be viewed under the "Literature Searching and Databases" section of the AWIC website.

An interested person (Dr. Breager) teaches an advanced trauma life support course at a medical school. He uses live pigs and dogs. The animals are euthanised at the end of the course. Although this has been the standard approach, Dr. Breager would like to know if there are any non-animal models or other alternatives.

Using the refined two-step approach, AWIC staff selects terms that would reduce and refine first, and terms for possible replacements last. Then terms are organised into a search strategy using all the information about searching presented above. Appropriate databases are chosen and the strategy is executed.

Databases selected

The databases are chosen because they have medical or veterinary resources, and technology citations from national and international sources. There is no limit to the years covered.

AGRICOLA	file 10
MEDLINE	file 155
NTIS	file 6
FEDRIP	file 266
BIOSIS	file 5
PASCAL	file 144
EMBASE	file 73
INSPEC	file 2

EI COMPENDEX file 8
JAPANESE SCIENCE AND TECHNOLOGY file 94

Search strategy
The results of the search strategy developed for Breager and executed in the above databases are as follows (tab. 1). Note: Sets 1, 2 and 3 contain terms and logic operators to retrieve citations that potentially offer reduction or refinement. S4 terminology is to retrieve replacement information. "RD" means duplicates were removed. In S6, the citations retrieved are to be limited to those with terms in the title, descriptor and identifier fields.

Some of the titles of citations retrieved:
Animal models
- Use of the ferret as a model for pediatric endotracheal intubation training. (file 10)
- Swine and dynamic ultrasound models for trauma ultrasound testing of surgical residents. (file 155)
- Ocular trauma modeling (file 73)
- Ultrasound training during ATLS: An early start for surgical interns. (file 5)
- *Battlefield Biomedical Technologies* (file 2)
- Removal of corneal foreign bodies: an instructional model (file 144)

Non-animal models
- Practicing procedures on the recently dead. (file 155)
- Virtual reality, robotics, and other wizardry in 21st century trauma care. (file 155)
- Paediatric resuscitation manikins. (file 155)
- Animal cadaveric models for advanced trauma life support training. (file 155)
- "Full scale" simulation in practical emergency medicine conception as represented by the Wurzburg anaesthesia and emergency medicine simulator. (file 73)
- MEDSIMM: Computer-simulated training for pediatric advanced life support (a work in progress). (file 5)
- *Medical Simulation for Trauma Management*. (file 6)
- *Enhanced Training using the Life Support for Trauma and Transport*. (file 144)
- Simulation and modeling of penetrating wounds from small arms. (file 2)

From the sampling of the results above, it seems likely that Dr. Breager will certainly have the opportunity to totally revise his training protocol! In fact it was interesting that a recent National Public Radio programme reported that a major university in the New York City area has completely changed their trauma training to non-animal technologies.

One last part of the process that should be addressed is the review of a search such as that above. If you are in the position of reviewer, look carefully at the following: the databases chosen, how the terms were selected and used, and how the Boolean logic was applied. If you do not perform each stage correctly, you can very easily come up with no information. A word of caution, if one does not perform each of the stages of the alternative search with intelligence and care, you can easily have zero hits returned. With the incredible volume of scientific information that has been carefully place in searchable database, zero hits are possibly a disservice to the scientific community and to the research animals.

The author hopes that the information included in this paper will provide information, knowledge, approaches, tips and tricks to formulate and execute an adequate and productive search for the 3Rs.

References

Jensen, D. (2005). *Workshop and Instructions for Alternatives Literature Searching*. Beltsville Maryland: ARS, NAL Animal Welfare Information Center.

Russel, W. M. S. and Burch, R. L. (1959). *The Principles of Humane Experimental Technique*. London: Methuen.

U.S. Government Printing Office (1985). The Animal Welfare Act – As Amended. Known as the "Improved Standards for Laboratory Animal" *7 US Code §§2131-2159. Public Law 99-198 7Code of Federal Regulations2.22, 2.80, and 371.2 (g).*

Acknowledgment

I would like to thank the USDA's National Agricultural Library for their continued support of the AWIC programme. A special

Tab. 1: Results of the search strategy

Set	Term Searched	Items
S1	trauma or life(w)support or emergenc?(1n)medic? or ems or emst or atls or advanced(w)trauma(w)life(w)support	440 688
S2	train? or teach? or educat? or instruct? or tutor?	2 118 835
S3	dog or dogs or canine? or pig or pigs or swine or piglet? or ferret? or cat or cats or animal or animals	2 419 446
S4	alternative? or model? or simulat? or cadav? or carcas? or software or video? or interact? or digital? or virtual or mannequin? or manikin? or computer?	16 081 397
S5	(s1 (4N) s2) and s3	621
S6	RD (unique items)	574
S7	(s1 and s2 and s4)/ti,de,id	1 387
S8	RD (unique items)	1 150

thanks goes to the talented, hard working and creative AWIC staff: Kristina Adams, Tim Allen, Sandra Ball, Barbara Buchanan, Dr. Richard Crawford, Judith Ho, D'Anna Jensen, Daniel Scholfield, Cynthia Smith, and the student Tyson Heckert.

Correspondence to
Jean A. Larson, M.A. B.A., A.A.S,
Coordinator
Animal Welfare Information Center
National Agricultural Library
10301 Baltimore Ave.
Beltsville, MD, 20705
Personal e-mail: jlarson@nal.usda.gov
e-mail: awic@nal.usda.gov
http://www.nal.usda.gov/awic

Considering Animal Alternatives and Welfare *via* a Comprehensive Search of the Scientific Literature

Mary W. Wood and Lynette A. Hart
UC Center for Animal Alternatives, School of Veterinary Medicine, University of California, Davis, USA

Summary
Searching the bibliographic databases for alternatives for the use of animals in research, teaching or testing, and considering alternatives to potentially painful or stressful procedures are regulatory requirements common to the US and member states of the EU, as well as several other countries. Effective searches are those completed in a variety of databases using a sophisticated and well considered strategy. Efficient searches are facilitated by the development and use of new search grids, practically organised by type of research, proposed animal model and research topic.

Keywords: alternatives, animal welfare, compliance, database

Introduction

The search for alternatives to comply with the related animal welfare regulations might seem to be a straightforward endeavour, yet resistance and scepticism continue from many scientists. Considering the Three Rs (replacement, reduction and refinement) and thus alternative experimental methods is now required by legislation in many countries. In the United States, the alternatives search requirement is included in the Animal Welfare Act in order to assure the general public that no animal used in research, teaching or testing will experience unnecessary pain or distress (United States Animal Welfare Act, 2002). In Australia, New Zealand and in the member states of the European Union, legislation requires that scientists provide evidence that the use of animals is justified and that they assess whether less stressful or painful procedures are available (Wood, 2005). While simply meeting the requirements for compliance is not difficult (a list of databases searched, the dates searched, the search terms used, and a discussion of what was found), giving thoughtful consideration to the breadth of alternatives and actually locating relevant information is uncommon. However, if a comprehensive search is performed, by searching all of the scientific literature published in a specific area of study, the information retrieved will be much more relevant, and proper attention to animal welfare is assured.

In order to identify and consider potential new methods and procedures, it is necessary to regularly search the published literature. Improved research methods may suggest alternative procedures that can reduce pain or distress, or even replace animal experiments. Already familiar with locating scientific material for their research, scientists can readily learn to expand their literature searches to include additional databases. At the same time, the concept of alternatives and the Three Rs objectives can be more clearly defined and explored. Expanding the search to include new ideas and technologies, as well as aspects of husbandry and care such as housing, blood collection, analgesia, anaesthesia and humane endpoints, will improve the utility of the results and facilitate compliance. New web-based resources located at the UC Center for Animal Alternatives main website, http://www.vetmed.ucdavis.edu/Animal_Alternatives/main.htm (Wood et al., 2005a) are designed to assist with a comprehensive search and to prompt the user to identify and consider additional concepts and databases.

Guide to Bibliographic Databases for Alternatives Searching

Scientists may locate information on their research, as well as on alternatives, in a variety of formats and systems, including printed and online bibliographic databases, specialised services and information managers and organisational and societal websites and resources. Even with so many useful options, the most common path to compliance is to search in a single database: PubMed. While this approach is clear-cut, identifying useful and relevant information is not a simple task. The principal barrier to effective searching and meaningful compliance is the misunderstanding of bibliographic searching (Grune, 2004). Together with an inconsistent interpretation of the Three Rs, searching for and locating quality information is a challenge. Each bibliographic database indexes a unique set of journals, but too often researchers use databases that do not cover their specific topics. Starting with an appropriate database that covers the type of information being sought is a first essential step towards conducting an effective search that can yield useful information to enhance animal welfare (Hart, 2005). In order to help researchers assess which databases may be appropriate when searching for particular types of information, we have developed a new web-based resource: *Guide to Bibliographic Databases for Alternatives Searching: Database Approach to a*

Search (Wood et al., 2005b). It is a simply presented grid arranged by both Animal Models and Topics. It includes live links to free databases that are accessible from around the world, including PubMed, AGRICOLA, PrimateLit, FishBase, TOXNET, ILAR Animal Models and others. Proprietary databases are listed as well, providing live links for affiliates to the University of California holdings. While only available to those with subscriptions, the listing of proprietary databases acts as an educational tool by increasing awareness of specific database availability. These are also arranged by Animal Model and Topic, and include PsycInfo, Zoological Record, Web of Science, BIOSIS, CAB and others. Depending on the institutional contract, scientists may be able to download articles in full-text.

Grids arranged by animal models and topics

This web-based *Guide to Bibliographic Databases* assists both scientists and members of animal care and use committees in determining whether a particular selected database is reasonable. It also works to inform on the availability of additional potentially worthwhile databases.

The list of journals indexed in any database is limited, and some journal titles are only indexed selectively. Published scientific information is so vast that each database must identify its niche and focus on a specialised area. The use of animals in science is such that it crosses a variety of research areas, and the Three Rs, especially, must be considered from many perspectives. Biomedical scientists commonly assume that the PubMed database is exhaustive, indexing every relevant scientific journal, cover to cover. While extensive, PubMed is not all-inclusive. The collection of journals in PubMed is oriented primarily on biomedical research and does not address the veterinary and animal science journals well, for example. A scientist proposing to use a bird or fish model may miss most of the published literature if PubMed is the only database searched; other databases to consider include BIOSIS and Aquatic Sciences and Fisheries Abstracts. AGRICOLA is an excellent database for animal science and veterinary literature, but focuses on US publications; CAB, with its European coverage, should also be searched to ensure broad consideration of this topic. Searching in multiple databases allows retrieval of different sources of information in order to realise a more comprehensive search.

Effective consideration of the Three Rs also requires using detailed search concepts, from specific research areas such as toxicity, tumour models and disease studies to more general issues such as analgesia, endpoints, imaging and euthanasia. By using the appropriate terms and resources, particularly useful material may be found, providing information that is otherwise very difficult to isolate. The grids in the *Guide to Bibliographic Databases* list common reoccurring topics, while providing links to specific resources, including books, bibliographies and specialised databases. For example, AWI Comfortable Quarters is valuable for housing questions, ILAR and AltWeb provide access to elusive information on humane endpoints, and ANZCCART Fact Sheets publish useful and very practical information on such things as managing laboratory animal pain. Topics such as blood collection, husbandry, identification, and behaviour, critical to the implementation of the Three Rs, are also topics that are frequently overlooked. Specialised publications and indexes, resources created by organisations and associations, have proven very valuable in the comprehensive search for alternatives.

Responsible literature searching

When animal welfare legislation requires the researcher and animal care oversight committees to consider alternatives and the Three Rs, a comprehensive and thorough alternatives search is expected. Supporting scientists' efforts in this area, and as part of a larger effort, a concise table of available databases and resources was recently developed, titled *Responsible Literature Searching* (Wood et al., 2005c). Organised by headings, "Research, Teaching and Testing," informational resources are listed and linked. The table is further divided into three sections: "Free Bibliographic Resources, Proprietary Bibliographic Databases and Free Governmental, Regulatory, and Organisational Databases". For each database, there is also a link to additional information about that database, allowing the user to learn more about the resource in order to determine whether it may be of relevance and should be included in the search.

If one is searching in the toxicology and testing area, there are several databases listed under the free databases section (i.e. Toxnet, EPA-EcoTox), a couple more listed as proprietary databases (i.e. ASFA and RTECS) and several more free resources under governmental, regulatory and organisational databases (i.e., ICCVAM, ECVAM, INVITTOX).

Similarly, under the topic of teaching resources, the user is provided with links to free databases, such as NORINA and AVAR, and to subscription-based databases, such as ERIC, CAB, and Zoological Record. At any time, the associated link will provide access to additional information on that particular resource, including its relevance to the proposed search.

Conclusions

The amount of scientific information currently being published and made available is overwhelming. If scientists are to thoroughly understand a proposed research study, or reliably decide on one method over another, they need help in managing the information. Their hope of finding just what they want when they need it requires some organisational assistance. The resources discussed here and found on the *UC Center for Animal Alternatives* website will help with that endeavour.

References

Grune, B., Fallon, M. Howard, C. et al. (2004). Report and recommendations of the international workshop "Retrieval approaches for information on alternative methods to animal experiments". *ALTEX 21*, 115-127.

Hart, L. A., Weng, H-Y. and Wood, M. W. (2005). Effective searching of the scientific literature for alternatives: search grids for appropriate databases. *Animal Welfare 14*, 287-289.

United States Animal Welfare Act 7USC 2131-2159, adopted 1966, amended 2002. Website: http://www.aphis.usda.gov/ac/awa.html.

Wood, M. W. and Hart, L. A. (2005a). UC Center for Animal Alternatives. University of California, Davis, School of Veterinary Medicine. Website: http://www.vetmed.ucdavis.edu/Animal_Alternatives/main.htm.

Wood, M. W., Hart, L. A. and Weng, H-Y. (2005c). Responsible Literature Searching. University of California, Davis, School of Veterinary Medicine. Website: http://www.vetmed.ucdavis.edu/Animal_Alternatives/responsible.html.

Wood, M. W., Hart, L. A. and Weng, H-Y. (2005). Responsible Literature Searching: Policies and Regulations. University of California, Davis, School of Veterinary Medicine. Website: http://www.vetmed.ucdavis.edu/Animal_Alternatives/policies&regs.html.

Wood, M. W., Weng, H-Y. and Hart, L. A. (2005b). Guide to Bibliographic Databases for Alternatives Searching: Database Approach to a Search. University of California, Davis, School of Veterinary Medicine. Website: http://www.vetmed.ucdavis.edu/Animal_Alternatives/databaseapproach.html.

Acknowledgements

The extensive electronic library holdings of the University of California make it possible to explore, identify and distinguish the unique features of the resources presented in these search grids. The authors appreciate Dr. Barbara Grune and the German Federal Institute for Risk Assessment (BfR) for their support and collaboration.

Correspondence to

Mary W. Wood, Librarian
UC Center for Animal Alternatives
School of Veterinary Medicine
University of California
Davis, CA 95616, USA
e-mail: mwwood@ucdavis.edu

Workshop 4.3
Search strategies – user requirements

Overview of the Regulatory Requirements for the Consideration of Alternatives

Jodie A. Kulpa-Eddy
United States Department of Agriculture (USDA)-Animal and Plant Health, Inspection Service (APHIS)-Animal Care, Riverdale, USA

Summary
This paper provides a brief history of the development in the United States of the requirement for the principal investigator to consider alternatives to procedures likely to cause pain and/or distress in animals. Countries that require similar considerations are also mentioned.
Inspectors with the U.S. Department of Agriculture (USDA), responsible for enforcing the Animal Welfare Act (AWA), expressed concern regarding researchers' compliance with this regulation, and examples from recent inspection reports are provided.
A list of guidance documents and practical suggestions are provided, followed by a workshop discussion on methods of measuring the impact of the use of alternative measures, and ways to distribute information on alternatives.

Keywords: consideration of alternatives, regulation, compliance

History

The Animal Welfare Act was enacted in the United States in 1966 in response to concerns raised by the public that stolen pets were being sold and used in research facilities. The initial regulations concentrated on record-keeping and identification, to create a method whereby animals could be tracked through the system. Standards of care were written for only six groups of animals, and centred on housing and husbandry. There were no regulations pertaining to the conduct of research using animals.

A groundswell of concern from the public began in the early 1980's regarding animal welfare in research institutions. The University of Pennsylvania Head Injury Clinic served as an example. The clinic had been conducting brain trauma experiments on nonhuman primates (baboons). Information about the experiments was taken from the laboratory and distributed to the public. Some of the procedures being used were questioned, and a National Institutes of Health (NIH) committee determined the procedures were inadequate to prevent the animals from suffering serious pain. Funding for the project was subsequently discontinued, and the publicity from this incident fuelled the public's desire to improve conditions for animals in research and to provide additional oversight.

It was a delicate balancing act, as Senator Robert Dole recognised "experimentation and testing on animals has benefited our society by yielding medical breakthroughs that have aided the development of new knowledge, new drugs and better surgical techniques which have saved countless lives". He understood "the use of animals will need to continue for the time being until alternative methodologies which do not use animals or which reduce the numbers of animals used and reduce the pain they experience can be further developed. At the same time, we need to ensure the public that adequate safeguards are in place to prevent unnecessary abuses to animals and that everything that is reasonably possible is being done to decrease the pain that animals suffer during experimentation and testing." (Dole, 1985).

These amendments to the AWA had the support of animal welfare organisations as well as the American Physiological Society, representing the largest group of users of experimental animals. They were passed by Congress in October 1985. As a result, the regulations promulgated by USDA to enforce the Act were updated. The regulations are intended to ensure that all possible steps have been taken to reduce or eliminate as much pain and distress as possible. This includes assurances that alternatives were considered with respect to all painful procedures

(regardless of whether or not pain relieving drugs are used or withheld) and that the research activities are not unnecessarily duplicating previous experiments.

The U.S. was not the first country to enact these requirements. Norway, Japan and the Netherlands were early pioneers in this area, and there were other international initiatives underway in the mid-1980s. The Council of Europe, European Treaty Series No. 123 on the "European Convention for the Protection of Vertebrate Animals used for Experimental and Other Scientific Purposes" was developed in 1986. It also includes language requiring the consideration of non-animal methods (if reasonably and practicably available), or justification for the use of animals, requiring the minimum number needed and causing the least pain, suffering or lasting harm yet still achieving satisfactory results (Council of Europe, 1986).

Today several countries recognise the need to review the scientific and ethical aspects of animal experimentation. A list of legislation to this effect may be found at http://www.vetmed.ucdavis.edu/Animal_Alternatives/policies&regs.html.

What is an alternative?

The USDA, Animal and Plant Health Inspection Service (APHIS), Animal Care refers to alternative or alternative methods as "those that incorporate some aspect of replacement, reduction or refinement of animal use in pursuit of the minimisation of animal pain and distress consistent with the goals or the research. These include methods that use non-animal systems or less sentient animal species to partially or fully replace animals (for example, the use of an *in vitro* or insect model to replace a mammalian model), methods that reduce the number of animals to the minimum required to obtain scientifically valid data, and methods that refine animal use by lessening or eliminating pain or distress and, thereby, enhancing animal well-being. Potential alternatives that do not allow the attainment of the goals of the research are not, by definition, alternatives." (USDA Policy 12, 2000)

Regulatory concerns

In 1999, USDA conducted a survey by mail of the inspectors of research facilities to "assess their opinions about the effectiveness of USDA's current approach to ensuring the humane care and use of animals at research facilities." (USDA Survey, 2000) One of the concerns identified was the "search for alternatives", with an estimated 600 to 800 facilities having difficulty in this area, making it the most frequently cited noncompliant item. "Avoiding unnecessary duplication" ranked fourth, and involved around 250 facilities.

The specific concerns cited by the inspectors in this survey included:

- inadequate justification provided for the number of animals used in research
- uncertainty by research facility personnel regarding when animals are experiencing pain/distress (examples: seizures, vomiting)
- inadequate documentation of literature searches for alternatives to painful procedures

A review of the 2004 USDA database of noncompliant items at research facilities again shows an inadequate "consideration of alternatives" (9 Code of Federal Regulations, Section 2.31(d)(1)(ii)) as the most frequent citation, but with a smaller number of facilities involved (8% of 1800 research facility inspections, or approximately 150 facilities). The regulation requires principal investigators to "consider alternatives to potentially painful procedures and provide a written narrative". The most common underlying causes for these citations are: (1) an inspector finds animals in pain/distress but there is no consideration of alternatives in the protocol, and/or (2) key words in database search are not related to the painful/distressful procedure.

Some specific examples:

- The search for alternatives to painful procedures (e.g. cutting marginal ear vein) has not been addressed and documented. A search for alternatives to distressful procedures (e.g. weight loss, snuffles and loose stools as a result of immunosuppression) has not been conducted.
- The protocol describes retro-orbital bleeding and toe clipping for hamsters and Jirds. Animals used are categorised as "no discomfort", and no consideration of alternatives was made.
- The written narrative descriptions did not provide enough details to determine what methods and sources were used to determine that alternatives were not available. Protocols indicated only that faculty read current veterinary publications.
- The phrase "in our experience" was used to provide a justification for the potential pain and distress to animals in the study.
- The PI (Principal Investigator) failed to conduct an adequate search for alternatives as the key words used were too vague. "Sheep", "alternative" and "pain" were used at one point, and "sheep", "alternative", "method" at another time in an Agricola search. Neither set does much to identify adequate alternatives to a skull implant procedure in sheep.

Ensuring compliance

The consideration of alternatives, and any justification for not using them, is of critical importance when the animals are likely to experience pain and/or distress as a result of the experiment. The principal investigator (study director) is primarily responsible for providing these assurances and must have training and instruction in how to utilise information services.

The Institutional Animal Care and Use Committee (IACUC, the reviewing committee) provides oversight at the institutional level. They are responsible for ensuring the investigator has made a "good faith" effort to consider and incorporate alternatives. What are some warning signs this may not have been done?

- A statement with no supporting documentation
- A "Google" search
- A review of one database for an area of research that is broader than the journals covered by that database.
- An old narrative (one that has not been updated in three or more years)

The 2000 USDA Employee Survey included a list of successful innovations observed by inspectors in use at various research facilities. Suggestions on how to enhance the development and review of a "consideration of alternatives" narrative include:

- Hire staff or use consultants (internal or external) who are statisticians or biostatisticians
- Include in the facility's training programmes:
  - o Proper methods for recognising, evaluating, alleviating and reporting pain
  - o Animal behaviourist
- Well-designed protocol forms and templates that require three separate answers regarding the need to use animals, the use of a particular species, and the use of a given number of animals
- The search for alternatives should be part of the planning and design stage of the research protocol, rather than considered as a "required afterthought".
- Pair new researchers with experienced librarians to assist with literature searches.

Additional suggestions that principal investigators and IACUCs may want to consider include:

- Guidance documents provided by USDA-APHIS-Animal Care:
  - o Policy #11 "Painful/Distressful Procedures" (http://www.aphis.usda.gov/ac/policy/policy11.pdf)
  - o Policy #12 "Written Narrative for Alternatives to Painful Procedures" (http://www.aphis.usda.gov/ac/policy/policy12.pdf)
  - o "Research Facility Inspection Guide" (http://www.aphis.usda.gov/ac/researchguide.html)
- Attending a workshop given by the USDA-Agricultural Research Service-National Agricultural Library-Animal Welfare Information Center (AWIC) (http://www.nal.usda.gov/awic/awicworkshops/awicworkshops.htm) or similar information service
- Utilising the AWIC alternatives sheet as an addendum to the protocol template (http://www.nal.usda.gov/awic/alternatives/searches/altwksht.pdf)
- Utilising listserves as a way to share information, ask others in the field for guidance, published papers, best practices
- Review the accessibility of free databases to search for alternatives (Donnelly, 2004)
- Review the accessibility of free databases to search for ongoing and previous research projects to avoid the unnecessary duplication of research. The Current Research Information Service (CRIS) is the USDA's documentation and reporting system for ongoing agricultural, food and nutrition, and forestry research. CRIS contains over 30,000 descriptions of current, publicly-supported research projects of the USDA agencies, the State Agricultural Experiment Stations, the State land-grant colleges and universities, State schools of forestry, cooperating schools of veterinary medicine, and USDA grant recipients (http://cris.csrees.usda.gov/). The Computer Retrieval of Information on Scientific Projects (CRISP) is a biomedical database system containing information on research projects and programmes supported by the US Department of Health and Human Services (http:// crisp.cit.nih.gov/).

Conclusions

The requirement to consider alternatives to painful procedures is a performance standard, not an engineering standard. It is understood that there are several means by which one may achieve the intended results. The universality of Russell and Burch's (1959) concept of replacement, reduction and refinement serves us well in this regard: to design research procedures that limit discomfort and pain to animals to that which is unavoidable for the conduct of scientifically valuable research.

References

Council of Europe (1986). European Convention for the Protection of Vertebrate Animals Used for Experimental and Other Scientific Purposes (ETS 123). Strasbourg, France: Council of Europe.

Dole, R. Senator (Kansas) (1985). "Improved Standards for Laboratory Animals Act," Congressional Record (4 June 1985) Daily ed. S14161.

Dole, R. Senator (Kansas) (1985). "Amendment No. 904," Congressional Record (25 October 1985) Daily ed. S29155.

Donnelly, Thomas M. (2004). Free Web Resources on Alternatives. Lab Animal 3, 46-48. http://www.labanimal. com/laban/ journal/v33/n3/full/laban0304-46.html.

Russell, W. M. S. and Burch, R. L. (1959). The principles of Humane Experimental Technique. London: Methuen.

USDA-APHIS [US Department of Agriculture-Animal and Plant Health Inspection Service]. 1989. 9 CFR Parts 1, 2 and 3: Animal Welfare; Final Rules. Federal Register 54(168):36123-36163(1989).

USDA-APHIS-AC [US Department of Agriculture-Animal and Plant Health Inspection Service-Animal Care]. 2000. Animal Care Survey: USDA Employee Survey on the Effectiveness of IACUC Regulations. April 2000.

USDA-APHIS-AC [US Department of Agriculture-Animal and Plant Health Inspection Service-Animal Care]. 2000. Policy #12. Animal Care Policy Manual, August 26, 2002.

Correspondence to

Jodie A. Kulpa-Eddy, D.V.M.
Staff Veterinarian, United States Department of Agriculture (USDA)-Animal and Plant Health, Inspection Service (APHIS)-Animal Care
4700 River Road, Unit 84, Riverdale, MD 20737-1234, USA
e-mail: jodie.a.kulpa-eddy@aphis.usda.gov

Theme 5
Safety testing, validation and risk assess

Session 5.01
Strategies for using non-animal methods in relation to chemicals legislation (HPV, REACH, ECVAM-Session)

REACH and CEFIC's Conception of a Feasible, Information and Priority Based Approach

Wolf-Rüdiger Bias
BASF Aktiengesellschaft, Ludwigshafen, Germany

Summary
The European Commission's proposal for a new Chemicals Legislation, REACH, presented on 29 October 2003, is currently under strong discussion in the processes of the first readings in the EU-Council and in parallel in the EU-Parliament. The legislative proposal is highly complex and sets out formal, volume dependent information requirements for the registration of substances. It calls for submission of extensive documentation of all information underlying a safety assessment. CEFIC proposes a leaner, more flexible approach. The prerequisites for a more efficient system shall be explored.

Keywords: REACH, chemicals assessment, pre-registration, prioritisation, exposure-related information

Introduction

REACH, the new European chemicals legislation, is in its fifth consecutive year of construction. The discussion of its main features has not yet reached a state that gives confidence that the act can enter into force by 2007, as currently envisaged by Council, Commission and Parliament of the European Union. What are the reasons for this controversial discussion? The most conflicting views seem to arise from matching the objectives of protecting man and the environment with the objective of competitiveness. All goals set out in the White Paper on 13 February 2001 on the Strategy of a future Chemicals Policy were welcomed by industry, especially since the environmental objectives were balanced with the economic one of "maintaining and improving the competitiveness of the chemical industry of the EU". This was also in line with the Lisbon Goal on competitiveness of 24 March 2000, namely to "make the EU the most competitive, dynamic and knowledge-based economic region of the world". However, if both goals, as formulated, are supported, a different approach must evidently be taken.

The new chemicals legislation is supposed to consolidate current chemicals legislation, adding consistency and efficiency in working with it, according to the conclusions of the Informal Environment Council held at Chester under the UK Presidency in 1998: "Environment Ministers in June 1999 called on the Commission to consider, *inter alia*, measures that provide an efficient and integrated design of the various legal instruments for chemicals; place the main responsibility on industry for generating and assessing data; provide a more flexible approach to risk assessment with the aim of targeting assessments; and establish effective risk management strategies for certain chemicals that may cause threats of serious or irreversible damage to human health or the environment as a result of their inherent properties by giving appropriate weight to their use pattern and the possibility of exposure" (Chester, 1999). In pursuit of its assignment, the Commission chose to focus on the following major pieces of European chemicals legislation:

(i) Directive 67/548 on classification, packaging and labelling of dangerous substances; (ii) Directive 88/379 on classification, packaging and labelling of dangerous preparations; (iii)

Directive 76/769 on the marketing and use of certain dangerous substances and preparations; and (iv) Regulation 793/93, the Existing Substances Regulation.

Outline and assessment of the legislative draft REACH

While the quoted laws that REACH is to absorb do not cover the entire area of EU chemicals legislation, they nevertheless encompass the universe of commercially handled substances. As proposed in the REACH draft, the steps evaluation and authorisation could be applied to any chemical, either on its own or as part of a preparation or an article manufactured in or imported into the EU. The registration is foreseen for all industrial chemicals exceeding a level of one ton per year per manufacturer or importer (fig. 1).

The registration would, despite some specific requirements and exemptions for certain classes of chemicals and certain uses, have to be performed for roughly 30,000 chemicals. It is therefore crucial to define the scope, the tasks, and the processes so precisely as to leave no doubt with the registrants about their duties and their freedom to act. There are two reasons to do so. First, because the entire registration has to be completed within 11 years (timeline, s. fig. 2), and second, because work that is not necessary for deciding on risk management measures, and, especially, duplication of vertebrate studies, must be avoided.

Due to the fact that only the four major pieces of European chemicals legislation shall become obsolete when REACH enters into force, the scope of REACH cannot be as clear cut as requested by the council. It has to take into consideration the obligations tied to many remaining laws regulating chemicals for specific applications. This requires excellent knowledge of the adjacent legislation on behalf of the registrant as well as on the part of the competent authorities. It also adds to complexity when having to observe a multitude of different laws in interfacing areas and therefore calls for clear legibility of REACH.

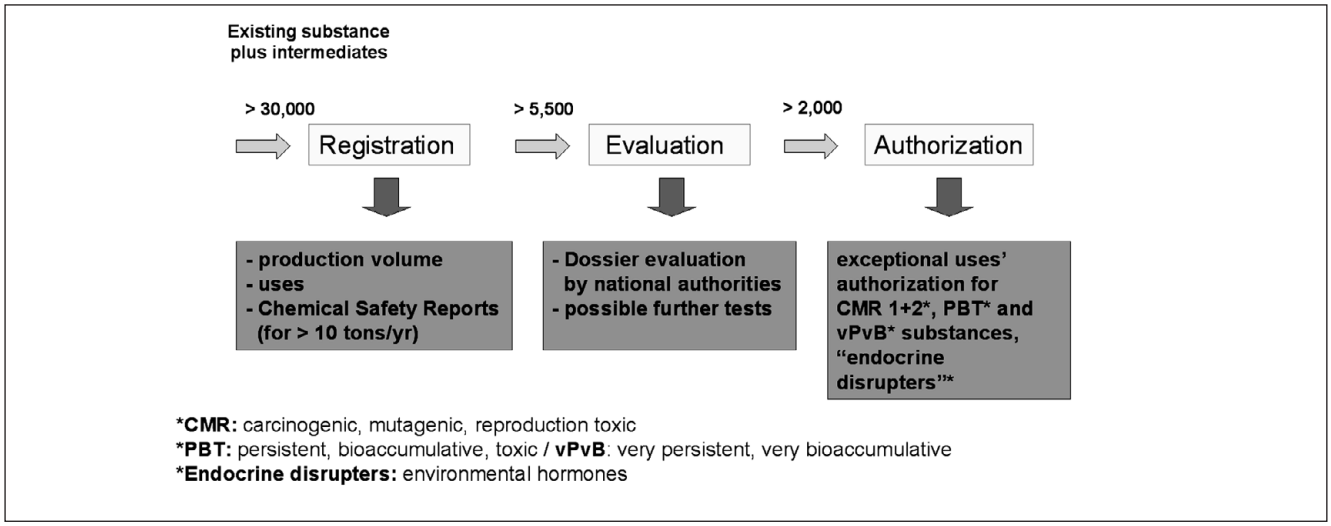


Fig. 1: Elements of REACH

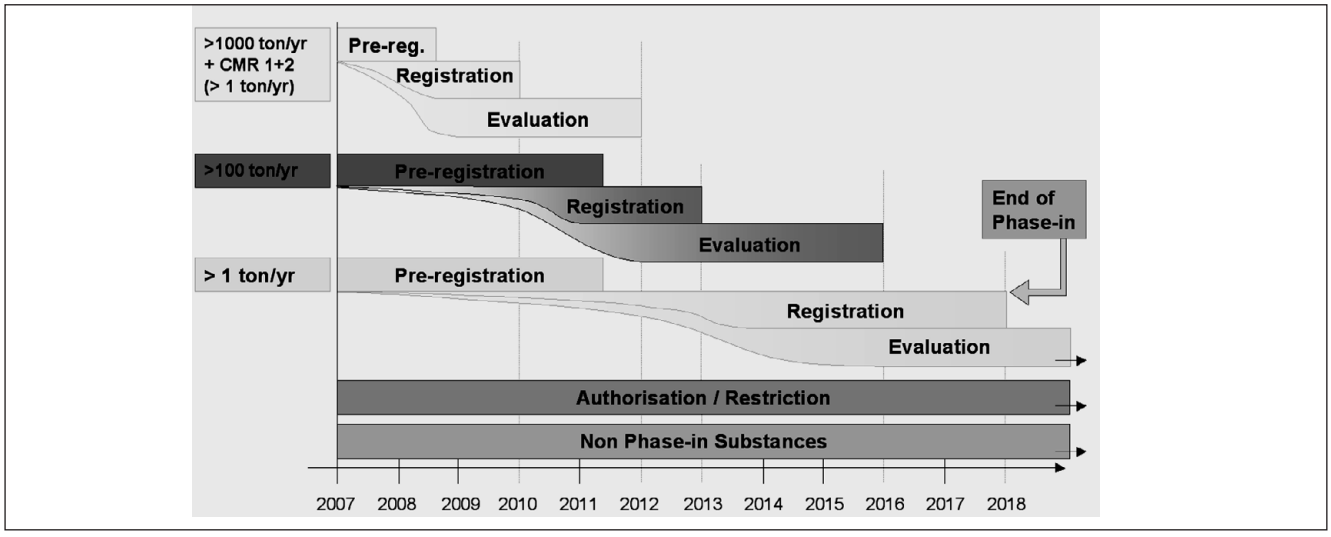


Fig. 2: REACH timeline – a process overview

Industry would favour legislation which, instead of offering workarounds for existing laws, would build upon a more systematic clean-out of the current chemicals legislation. For example, the usefulness of the Biocidal Products Directive 98/8/EC in the presence of REACH is questioned as much as those pieces of various legislation setting out information requirements on substances as a basis for an assessment. A clear, exposure-related approach, triggering the information needed for any safety assessment in REACH, could be the basis to many a legislation once generally accepted strategies on information requirements have been established.

The lack of a distinct tier for exposure-related information requirements is one of the major shortcomings of REACH. Instead, in overestimating the influence that the volume of a substance may exert on risk, a tonnage-related trigger has been established. Volume usually is a minor cofactor in determining exposure-regimes. Adhering to the volume-trigger has not only prompted numerous proposals by Industry to find compromises with the Commission. In effect, it leads to furnishing substances with uniform sets of information generated according to a formal standard as required under annexes V-VIII, rather than focusing on uses that lead to crucial exposure-regimes and which call for more information for their assessment. This way, not only resources are spread and spent evenly, but work on substances may also be allocated to the wrong point in time.

On the whole, it is also by far underestimated how difficult volume tracking within one company actually can be. Such figures need to be established by diligent tracing of constituents in products, their transfer to other business units, and adding up their total quantities while, at the same time, different points of entry within one legal entity have to be observed.

Regarding the registration of chemicals, the Commission has proposed a staggered approach for pre-registration and the submission of registration documents according to the volume manufactured or imported per registrant. While the timeline is highly ambitious, the current scheme with two deadlines for pre-registration has another disadvantage. It leads to uncertainty among the market participants as to whether substances are intended to be registered and available in the future. Moreover, it would lead to registration of one and the same substance at different points in time if one manufacturer/importer has to register early because of the large volume he manufactures/imports and another, who only deals with small volumes of the same substance, submits his vertebrate data at a later stage.

For the purpose of registering a substance, the registrant submits the required documents in the form of a technical dossier and additionally, for all substances above 10 tons/year, a Chemical Safety Report (CSR) to the Agency. The Agency will check the documents for completeness but not for quality, ask for supplements if parts are missing and pass on the dossiers to the competent authority of the relevant member state within 30 days. It is the task of the member state's competent authority then to examine any test proposals made by the registrant and to check for any dossier, whether it complies with the rules as set out, or to decide whether additional information is needed for the dossier to comply with the relevant information requirements. The registrants of a substance can either jointly submit a dossier or choose to register individually. As seen from work under the Existing Substances Regulation, a considerable effort is needed to coordinate cooperation. For low volume substances it is therefore expected that the latter will most likely be the prevailing situation. This will entail that each relevant member state competent authority will separately evaluate the dossiers of the manufacturer or importer resident to his state (fig. 3).

Such a process calls for improvement. Not only that the workflow between the central Agency and the member state's competent authorities is complicated. Evaluation by several competent authorities of different member states would be a duplication of work. There is also the likelihood that different competent authorities come to different conclusions on the proposals, justifications and assessments taken.

company A manufactures substance X in Germany → Evaluation by German authorities

company B manufactures same substance X in France → evaluation by French authorities

company C imports same substance X into Spain → evaluation by Spanish authorities

Consequence:
- double work with authorities*
- no build-up of competence
- avoidable costs

*In case companies A,B & C form a consortium only the authority of the resident registrant submitting will initially have to act

Fig. 3: Dossier-Evaluation: doubling the work in Europe

- one pre-registration step only
- maintain tonnage trigger for administration, but introduce more flexibility through:
  - prioritisation and
  - exposure-related information needs
- central agency responsible for evaluation and all subsequent steps

Fig. 4: Key elements of the CEFIC proposal

- one-step obligatory pre-registration, within 18 months to be sent to the Agency with the following information:
  - company name
  - volume band
  - substance name and CAS n°
  - availability of vertebrate studies
  - interest in joining consortia and sharing of own data

Fig. 5: Pre-registration

- Prioritisation procedure
  - prepare a Core Information Set for all substances above 10 tons and assemble appropriate available information for substances 1 – 10 tons/yr
  - identification of substances of potential high risk
  - Core Information Set for substances above 10 tons/yr together with the prioritisation to be sent to the Agency within 5 years after REACH has entered into force

Fig. 6: Prioritisation

Physico-chemical data	• Physical form (particle size) • Melting point • Boiling point • Vapour pressure • Partition coefficient octanol / water • Relative density • Explosiveness • Flammability • Water solubility • Flash point
Ecotoxicity	• Biodegradation • Acute toxicity (daphnia)
Toxicity	• Acute toxicity – relevant route • Skin irritation • Eye irritation • Skin sensitization (if no clues available) • Genetic toxicity bacterial test (Ames)
⇩ ⇩	• Classification & Labelling • Generic Exposure Information

Fig. 7: Core Information Set

In order to succeed, a series of further stipulations of REACH call for very carefully designed processes, which the proposal currently does not provide. The unsolved issues of maintaining the protection of confidential business information, of unrestricted ownership rights in studies, contractual antitrust agreements for consortia formation, and information flow through the supply chain give a perspective to which extent bureaucratic rules might be needed to cope with these issues, making REACH theoretically safe but non-sustainable by practical means.

Remedies for selected shortcomings of the current REACH proposal

CEFIC, the European Chemical Industry Council, has accompanied the development of REACH with numerous constructive proposals made in order to obtain a viable legislation. While drafting the proposal, the Commission took only few of them into consideration. If the present structure of REACH is to be maintained, CEFIC advocates substantial changes, some of which shall be outlined below (fig. 4).

Referring to the assessment of the current proposal above:

- Pre-registration should be a one-step obligatory affair, in order to allow simultaneous registration of the same substance regardless of the quantities manufactured/imported at one point in time (fig. 5) and to provide certainty about the availability of substances in the future. This would be the initial step to establish a new chemicals inventory within 18 months time.
- The tonnage trigger shall only be maintained as an administrative tool to divide the workload and to define the maximum in proportionate information requirements. However, two features are added to achieve a much more flexible approach: They consist of prioritisation (fig. 6) and exposure-driven information needs. For all substances above 10 tons/year a Core Information Set (fig. 7) shall be elaborated within a period of 5 years. On the basis of a potential for risk, as suggested by inherent properties like vapour pressure, dustiness and high toxicity as well as the use domain, substances will be prioritised by very simple arithmetic (fig. 8). An exposure analysis could not be established within such a time frame, which, together with hazard information, would allow addressing the risk. Therefore only the potential for risk shall be utilised.

The prioritisation shall also help to spread the workload more evenly. CEFIC proposes to adhere to registering high volume and cmr substances with the first lot, but with a time frame of 5 years. The next lot would encompass substances of 100-1000 tons/year as well as those of lower volume per year with a potential for elevated risk (fig. 9).

Finally, the remaining substances with a potential for lower risk shall be registered.

- For all substances, based on the available information and above 10 tons/year based on the Core Information Set, the use-related exposure should determine whether additional information is needed to the maximum of volume-relating annexes.

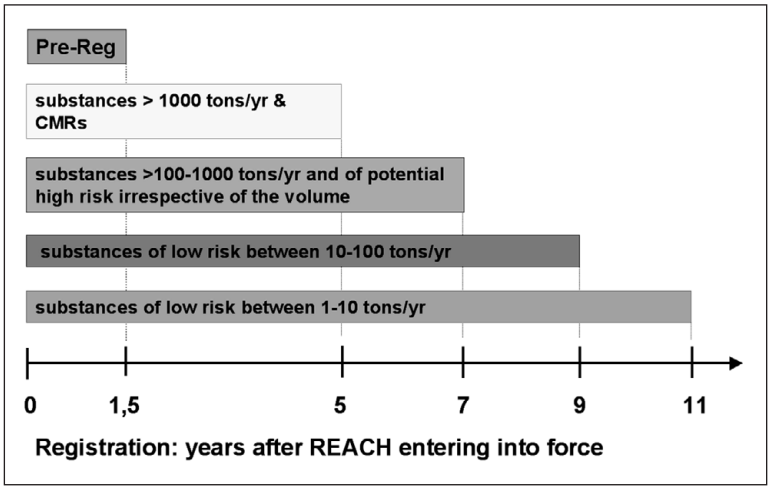


Fig. 8: Registration schedule

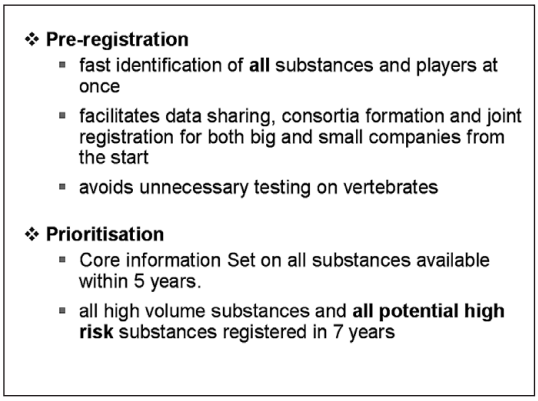


Fig. 9: Benefits: procedural

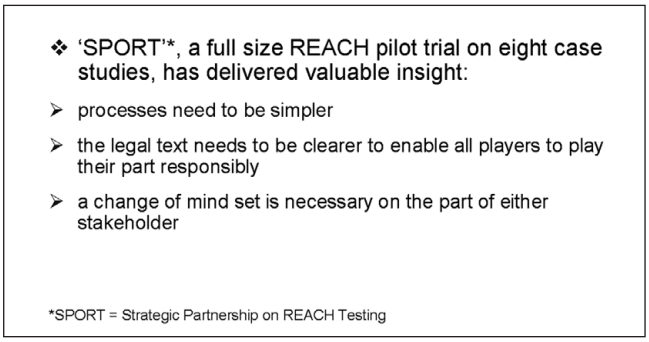


Fig. 10: Recommendations from 'Field Studies'

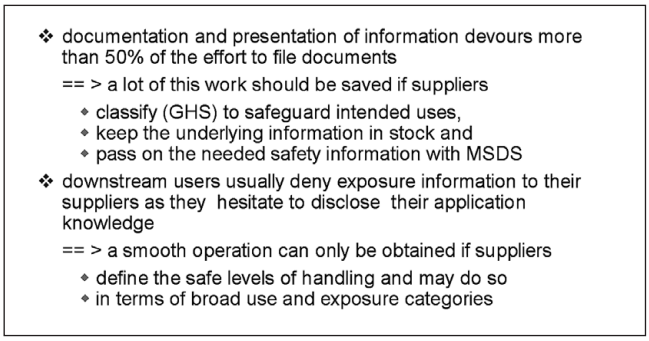


Fig. 11: Less complexity for better performance

- To warrant that evaluation is conducted following the same principles and rules throughout Europe, the responsibility for the evaluation shall reside with the Agency. This would entail an alleviated information flow, improved consistency, and would help avoid duplication of evaluation.

It is obvious that the proposed steps are not the only prerequisites for a viable REACH. Many more proposals have been made by CEFIC, which cannot be presented here.

- The recent pilot study on REACH called SPORT just completed in the middle of this year (fig. 10). SPORT, i.e. "Strategic Partnership on REACH Testing" was an Industry initiative jointly supported by the EU Commission, Member States and Industry. More than 150 recommendations were distilled from 8 projects conducted under the conditions of the current REACH proposal. The joint recommendations by the three partners called for simplification of processes to enable every stakeholder to contribute according to his skills and capacity. They furthermore set out that REACH needs to be understood by everyone so that everyone can live up to his/her obligations. Also, a uniform view was reached that exchange of data on structurally related compounds should be facilitated in order to enable grouping of substances and to maximise the use of information to avoid duplication of vertebrate studies.

Reduced complexity is the only way to avoid running into a high implementation deficit very quickly. A way forward to improve the performance further is outlined in figure 11 as a personal view. I sincerely hope that Commission, Council and Parliament now demonstrate their flexibility to set the controls for a lean but efficient REACH.

Correspondence to
Dr. Wolf-Rüdiger Bias
BASF Aktiengesellschaft
D-67056 Ludwigshafen
Germany
e-mail: ruediger.bias@basf-ag.de

Strategy for Minimising the Use of Animal Testing as Part of REACH

Uwe Lahl
Federal Ministry for the Environment, Nature Conservation and Reactor Safety, Bonn, Germany

Summary
This paper outlines the current state of decision-making in the reform of EU chemicals law. REACH will mean that we will acquire extensive data on the properties of chemicals and existing exposures. This will make it possible to reduce the negative effects they have on humans and the environment. Although consumer protection as well as occupational health and safety are at the centre of the public discussion, the implications for environmental protection should not be neglected. Many dangerous environmental chemicals (e.g. endocrine-disrupting substances) contribute to the fact that individual animal populations are in decline regionally and even globally. REACH is thus also needed to better protect wildlife from environmental chemicals.
The aim to protect wildlife is at odds with the fact that it is not always possible to ascertain the risks associated with these chemicals without carrying out tests on animals. However, the Commission's proposal for a regulation identifies all the endeavours being made to limit the number of animal tests. OSOR will prevent duplicate or even multiple tests on animals. The use of historical data will also be permitted, provided that they are still valid. Alternative methods to replace animal testing (particularly in vitro methods) will also be used wherever available. Opening up risk analysis to methods that analyse structure/activity relationships (in silico methods, SAR, QSAR) creates further prospects for reducing the extent of animal testing in the medium term.
Finally, the time between now and the date REACH enters into force is a clearly defined time frame within which current research into alternative methods can be advanced to a state that will allow additional possibilities for replacing animal testing to be developed.

Keywords: REACH, EU chemicals legislation, animal testing, three Rs concept, minimisation strategy

Introduction

The discussion on animal testing has been running for many years and is highly controversial[1]. For example, people question the fundamental ethical and moral right to carry out experiments that (may) cause suffering, detriment and even death to animals[2]. This stance was adopted particularly with regard to animal testing in the area of pharmaceutical research, cosmetics and basic research. These are the areas in which the highest numbers of animal tests are conducted (tab. 1).

Fundamental arguments and ethical reasons are cited in standpoints that justify animal testing[3]. For example, it is said that animal testing leads to extremely important additional knowledge. Furthermore, it is seen as an ethically acceptable consideration that the suffering of animals should be tolerated in order to prevent human suffering. In recent years the discussion has split into different camps. However, all the groupings agree that animal testing should be reduced to the absolute minimum. For that reason, animal tests are subject to a permit or notification procedure under which the justification for the test has to be submitted[4]. The competent authority bases its decision on a prescribed set of criteria[5].

For regulatory practice, the ethical consideration presents itself in a very particular way. Here, the benefit of laboratory tests on animals is not simply that humans are spared suffering and damage by the way chemicals are marketed today. Animal tests are also used to identify the harmful effects of substances on the environment. Avoiding or reducing this impact is the ultimate aim of REACH. It is to be achieved by testing and by management measures following the evaluation of the substances. Only in this way will it be possible to avoid or reduce suffering and harm to animals in the wild, which can otherwise, in extreme cases, go so far as to cause the extinction of individual species.

[1] On this topic see: http://www.verbraucherschutzministerium.de/ and http://www.tierschutz.de/ and http://www.bmt-tierschutz.dsn.de/ and http://www.tierschutzbund.de/ and http://www.ihep.jrc.it/ and http://www.ebra.org/ and http://www.bfr.bund.de/cd/1433
[2] Wolf U.: Tierversuche und angebliche moralische Konflikte. http://www.tierversuchsgegener.org/Tierversuche/ursula.wolf.html und Wolf J.-C.: Warum Tierversuche moralisch unzulässig sind. http://www.tierversuchsgegner.org/Tierversuche/jean-claude.wolf.html and http://www.tierversuchsgegner.org/Tierversuche/hartinger.html
[3] German Research Foundation: Tierversuche in der Forschung. VCH-Verlagsgesellschaft, 1993
[4] Animal Welfare Act. BGBl. I (Official Gazette) p. 1105, 1818 and Allgemeine Verwaltungsvorschrift zur Durchführung des Tierschutzgesetzes of 9 February 2000 (Secondary legislation implementing the Animal Welfare Act)
[5] Under the Animal Welfare Act, animal tests that are required by law (e.g. under pharmaceutical or chemicals legislation) are merely notifiable.

Tab. 1: Number of animals used in Germany in 2002 by purpose of test*

Basic biological research	Development & quality assurance of pharmaceuticals for human and veterinary use	Toxicological studies, safety testing	Diagnosing illnesses	Other purposes
826,729	854,078	207,511	50,700	273,358

*Federal Ministry of Consumer Protection, Food and Agriculture: Tierschutzbericht 2003

REACH

REACH stands for the Registration, Evaluation and Authorisation of Chemicals[6]. Without going into the details of the EU Commission's proposed regulation, we shall describe just those aspects that are of greatest importance for the topic under consideration here (animal testing). For more extensive treatment, please refer to the relevant literature[7,8,9,10].

Approximately 30,000 existing chemicals with sales volumes in excess of one tonne per annum are on the market in Europe. Under REACH, these substances will be subject to registration with a central agency. In connection with the registration, the applicant will have to submit a set of data describing, amongst other things, the intrinsic effects of the substance. The effect data are supplied by the manufacturer of the substance. If the manufacturer or the government agencies involved identify any "risks" on the basis of the hazards already ascertained, the substances will undergo more intensive evaluation. For particularly critical substances, the manufacturer will have to provide evidence of their safe use and apply for their authorisation. He will only receive authorisation if the governmental agency is convinced by the evidence provided and approves the use.

The proposal for REACH is currently being discussed in the European Parliament and Council. The EU Parliament is scheduled to complete the first reading in autumn this year and it is envisaged that a common position will be agreed in the Council under the British presidency. It is to be expected that a process of mutual consultation between the Commission, Council and Parliament will take place at the beginning of 2006. According to this timetable, REACH will then enter into force in 2007. Setting up the working structures needed and implementing the necessary secondary legislation will occupy the whole of 2008, so that REACH will necessitate data acquisition on a significant scale in 2010 at the earliest (*in other words, in over five years time*).

[6] http://europa.eu.int/eur-lex/de/com/pdf/2003/com2003_0644de.html
[7] http://www.umweltbundesamt.de/reach/
[8] http://www.bmu.de/chemikalien/doc/6486.php
[9] http://www.bmu.de/chemikalien/doc/6073.php
[10] http://www.bmu.de/chemikalien/aktuell/doc/35399.php

Tab. 2: Testing requirements under REACH as set out in Annexes V to VIII of the EU Commission's proposed regulation

> 1 t/a	> 10 t/a	> 100 t/a	> 1,000 t/a
Skin irritation	*In vivo* skin irritation test		
Eye irritation	*In vivo* eye irritation test		
Skin sensitisation			
Mutagenicity *In vitro* Ames-Test	*In vitro* mammalian gene mutation *In vitro* cytogenetics test		
	Acute toxicity		
	28-day test		Possibly long-term toxicity > 12 months
	Screening for toxicity related to reproduction	Toxicity related to reproduction for one animal species	Two generation reprotoxicity
Short-term toxicity for daphnia		Long-term test for daphnia	
	Growth-inhibitor test on algae		
	Short-term toxicity for fish	Long-term toxicity for fish	
	Absorption/desorption screening test	Further absorption/ desorption tests	
		Accumulation in fish	
		Short-term test on earthworms	Long-term test on earthworms
			Long-term toxicity for invertebrate animals
			Long-term toxicity for birds
		Short-term toxicity for plants	Long-term toxicity for plants

The data will be collected over a 10 to 15 year period in a tiered procedure based on quantity thresholds and risk criteria.

REACH will entail animal testing

The tests to be conducted under REACH are laid down in Annexes V to VIII. If the Commission's proposal is accepted, no tests on animals will be necessary for the basic data set in the volume range of 1 to 10 tonnes per year, as shown in table 2.

Above the 10 tonne per year threshold, *in vivo* tests for skin and eye irritation and toxicity in fish must be performed, although only if prior *in vitro* tests suggest that the test animal's skin will *not* be damaged. The purpose of the tests is thus merely to confirm negative results. Toxicologists consider this additional back-up necessary. Furthermore, the results of animal tests from what is known as the "28-day test" must also be submitted. This test provides initial indications of the effect of repeated exposure to a substance. This test is essential to carry out risk assessment for health and safety at the workplace, for example. A screening test for reprotoxicity gives initial indications of the possible risk to reproduction.

For the purpose of human toxicology, additional animal tests are required above the 100 tonne per year mark. They include a sub-chronic toxicity test (90 days), a two-generation test and a teratogenicity test to investigate toxicity related to reproduction. In the area of ecotoxicology, long-term toxicity to fish is one of the areas studied.

Ultimately, for production volumes in excess of 1,000 tonnes per year, further toxicological tests are only necessary in isolated and justified cases, e.g. to clarify suspected carcinogenicity. The tests for ecotoxicity include a test on birds.

Above and beyond the animal tests listed in table 2, the competent authority can demand further tests, if appropriate, as part of the process of evaluating a chemical. This will be the case when the set of data submitted shows grounds to suspect harmful effects that need further clarification. This evaluation is carried out for chemicals that, on the basis of the data submitted, are classified as high-risk or suspect chemicals and that therefore had to be placed on a rolling plan.

Animal testing only where absolutely essential

Of course, it is important to analyse critically whether the above-mentioned animal tests are necessary in cases where industry does not have the required data already. The people with the political responsibility for this decision are therefore asking scientists whether it is possible to evaluate the risk of chemicals in terms of chronic eco- and human toxicity without animal testing – using existing alternative methods, for example. The panel of experts appointed by the Commission (CSTEE, now Scher) gave its opinion on this in an extensive report published last year[11]. The background to this report was a proposal by the British anti-vivisection organisation to end animal testing under REACH and use alternative methods instead[12,13]. The panel of experts came to the conclusion that the animal tests mentioned cannot currently be replaced and that not carrying them out would significantly diminish today's level of health and environmental protection from hazardous chemicals.

This verdict nevertheless does not release us from the obligation to make every conceivable effort to reduce the number of animal tests required. However, it would also not be appropriate for politicians to disregard the verdict.

The 3R principle

As early as 1959, scientists Russel and Burch proposed a way of consistently reducing the number of animal tests. The 3Rs principle – replacement, reduction, refinement – is an integral part of the Commission's proposed REACH regulation.

Reduction

It is possible to reduce the number of animals used for testing without diminishing the informative value of the test or its validity. Table 3 shows the successes that have been achieved by reducing the number of test animals, illustrated by the example of the LD_{50} test used to determine acute toxicity.

A further aspect in reducing the number of test animals is the non-performance of duplicate or even multiple tests. It was by no means unusual in the past to see the duplication of tests. One reason for this was that national agencies did not reciprocally recognise each other's tests, but another frequent reason was that tests of that kind are expensive and companies regarded them as their personal property, so that sharing data with other companies did not happen automatically.

[11] European Commission, Health & Consumer Protection Directorate: Opinion of the Scientific Committee on Toxicity, Ecotoxicity and Environment (CSTEE) on the BUAV-EDEA Report. Brussels 8 Jan. 2004; http://66.249.93.104/search?q=cache:vYwulqWfyXwJ:europa.eu.int/comm/health/ph_risk/committees/sct/documents/out217_en.pdf+BUAV-ECEAE&hl=de

[12] BUAV-ECEAE Report: "The way forward – Action to end animal toxicity testing" http://66.249.93.104/search?q=cache:m3sUmnMOu5QJ:www.ne.jp/asahi/kagaku/pico/eu/reach/eu/REACH_Part_3.pdf+BUAV-ECEAE&hl=de

[13] See also: http://www.reachnonanimaltests.org/

Tab. 3: Success in reducing the number of test animals using the example of one test

In the area of toxicology, 37% of all animals are currently used for LD_{50} tests	
1970's	150 animals per substance
1980's, following harmonisation by the OECD	45 animals
1990's, tiered testing strategy	15 animals
Current ECVAM/ICCVAM test	5 animals ??
Future: *in vitro* test	0 (no animal)

In recent years, the OECD in particular has rendered outstanding service in this area[14]. The principle of MAD (Mutual Acceptance of Data) has been achieved through standard test methods and GLP (Good Laboratory Practice). This has led to a significant reduction in the number of animal tests, because the system is used worldwide (i.e. also by non-OECD countries). REACH is based on GLP and MAD and has developed this set of regulations further. Thus, every possibility of reducing the number of animal tests is being exploited.

"Officially," i.e. according to statements issued by the Commission, the Commission's proposal for REACH contains an obligation to use existing data from previous animal tests. However, there are in fact some gaps in the regulations here. This means that under the Commission's proposal duplication of tests is theoretically possible. Germany has thus submitted suggestions for corresponding text amendments to close these gaps.

One way in which the number of tests carried out will be reduced is the fact that multiple registrations of a single substance will be avoided. The avoidance of duplicated animal tests will be achieved under OSOR (*O*ne *S*ubstance, *O*ne *R*egistration) by the obligation to submit one joint set of core data per tonnage band. For subsequent tonnage bands, use should be made of the substances already registered (through a Substance Information Exchange Forum (SIEF) per tonnage band). Obliging companies to jointly submit a core data set also facilitates a better overview of existing (test) data. Under the Commission's REACH proposal, all the animal tests listed in table 2 for the 100 tonne per year category are subject to a proviso. They are not to be carried out automatically but only if the data is required for risk assessment. To this end, the manufacturer of a particular chemical has to submit a proposal for the further tests. He also has the possibility at this point to make a case for why the data are not necessary. Only when the competent authority has reviewed this proposal as part of its obligatory evaluation of the dossier and has come to a conclusion, are the tests listed in table 2 carried out.

Under REACH there are three possibilities for waiving tests:
1. the possibility that is fundamentally always there of not carrying out tests if they are not scientifically necessary or technically feasible;
2. the general stipulations in Annex IX of the REACH Regulation for deviations from the standard testing programmes set out under Annexes V to VIII;
3. special conditions for waiving individual tests (e.g. as stipulated in Annex VI.6.6.1: the 28-day test does not need to be conducted if relevant human exposure can be ruled out).

Refinement

Refinement in this context means carrying out the testing in a way in which additional information on toxic properties is acquired without the use of additional animals. This in turn means that further studies are not required. In recent years it has, for example, become possible to replace acute tests on fish by a test on fish eggs or embryos[15]. By incorporating modern molecular biological methods ("-omics" techniques, "toxicogenomics") into classic tests, it is already possible to classify chemicals in a particular category and thus predict their charac-

[14] http://www.oecd.org/searchResult/0,2665,en_2649_201185_1_1_1_1_1,00.html

[15] Gies A. et. al.: Replacing vertebrate testing in regulatory ecotoxicology, Umweltbundesamt, Workshop Ecotoxicity – Applying the 3 Rs, www.bmu.de/files/chemikalien/downloads/application/pdf/reach_minimierung_tierversuche.pdf

Tab. 4: *In vitro* testing as an alternative to animal testing (according to the Federal Institute for Risk Assessment – BfR)

OECD test method	Validated and accepted	Under development	Planned
Skin absorption, *In vitro* method **OECD TG 428**, Accepted 13.04.2004	X		
Skin corrosion, *in vitro*, Rat skin model **OECD TG 430,** Accepted 13.04.2004	X		
Skin corrosion, *in vitro*, Human skin model **OECD TG 431,** Accepted 13.04.2004	X		
***In vitro* phototoxicity test,** 3T3 NRU phototoxicity test **OECD TG 432,** Accepted 13.04.2004	X		
Acute inhalation toxicity Fixed Concentration Procedure – FCP **OECD TG 433,** 2nd draft 08.12.2004		x	
Acute dermal toxicity, Fixed Dose Procedure – FDP **OECD TG 434,** 1st draft 08.12.2004			x
In vitro micronucleus test **OECD TG 487.** 1st draft 14.06.2004			x
***In vitro* skin irritation,** Skin irritation *in vitro* ECVAM validation study 2003-2005		x	
Eye irritation, Acute eye irritation/corrosion **OECD TG 405,** ECVAM/ICCVAM co-operation project		x	
Chemical carcinogenicity , SHE cell transformation assay **OECD Draft TG,** ECVAM validation proposed		x	

teristic toxicological profile or a significant property, such as carcinogenicity. The use of these molecular biological methods has great potential. Their development should therefore be promoted more intensively in the future.

Another important aspect is making use of prior information (e.g. from screening assays, information about probable effect mechanisms) to carry out any necessary tests on vertebrates in such a way that the maximum amount of information for the relevant regulatory issue is obtained using the minimum number of animals (see toxicogenomics above and *in vitro* methods below).

This requires clear criteria and rules that have been agreed by all parties involved that set out the conditions under which tests are essential or can be waived. During the implementation phase of REACH[16], intensive discussions that incorporate a broad range of scientific and regulatory expertise will be necessary.

Replacement

When it comes to replacement, two different fields of action can be distinguished:

- the development of *in vitro* methods and models,
- the development of *in silico* methods and models.

From the regulatory point of view, it must be stressed that the "intelligent combination" of *in vitro* methods can help to minimise the number of definitive tests necessary (which as a rule are animal tests) and to maximise the information yield from those animal tests that are necessary. The successful implementation of these two aspects (1: marked reduction in the number of animal tests required and 2: the prevention, as far as possible, of false negative results) will probably always require the combination of several *in vitro* methods. Here it must be said that the cost involved in "intelligent" combinations of testing strategies should not be underestimated.

One obstacle to these methods becoming more widespread is the question of recognition (validation). This recognition often takes place in international bodies such as the OECD. The following table shows the current state of progress and a prognosis for the period up to 2010.

Table 4 shows that *in vitro* tests for many relevant endpoints still have to be developed or validated. It follows therefore that the results achieved are not yet satisfactory. Unfortunately, currently available *in vitro* test methods (that do not use animals) cannot determine to a fully satisfactory level how dangerous chemicals are. This also emerges from a report published by ECVAM[17]. It seems therefore that efforts in this direction need to be intensified. In my opinion, the availability of research funds should not be a problem here.

"*In silico* toxicology" (computer-aided analysis of effects and their interactions) is a relatively young discipline, but it already has a broad range of applications and methods[18]. It searches for connections between effects and chemical structures in order to make prognoses.

The models can to some extent be used for different effect endpoints. Table 5 shows a selection of the *in silico* models available today for endpoints that are significant for human toxicology and ecotoxicology. The decisive point for the issue we are looking at here (animal testing) is whether the methods mentioned can also produce sound results. To try and determine this, the European Centre for Ecotoxicology and Toxicology of Chemicals carried out a study, the results of which are summarised in table 5 (32).

The study shows that validated models are available for particular endpoints, especially in the area of ecotoxicology, and that they provide sound results. But this is not the case for the issue under discussion here, i.e. the long-term test for chronic toxicity to humans. However, the available models do already provide important additional information that can be used to back-up evaluations.

The applicability of an *in silico* model is highly dependent on the chemical structure of a substance. In other words, all models are not equally well suited for all chemicals. To improve the applicability of *in silico* methods we must therefore establish which methods are suitable for which chemicals. That is the only way to establish whether an *in silico* method is applicable and what level of uncertainty is associated with the prediction.

It is not yet possible to significantly reduce the number of animal tests on the basis of the knowledge described above. The majority of toxic effects of substances are currently still difficult to predict. It would therefore seem essential to step up efforts in the area of *in silico* methods, too. In my opinion, the availability of research funds should similarly not be a problem here. It

[16] REACH Implementing Process (RIPs): http://ecb.jrc.it/REACH/RIP_PROJECTS/

[17] Worth A.P., Balls M. (Eds.): Alternative (Non-animal) Methods for Chemical Testing. A Report prepared by EVCAM, ATLA 30, 2002

[18] http://www.ndsu.nodak.edu/qsar_soc/resource/software.htm

[19] http://www.bgvv.de/cm/232/alternative_testverfahren_und_intelligente_teststrategien_position_aus_sicht_der_wissenschaft.pdf

Tab. 5: Applicability of *in silico* models for endpoints relevant to ecotoxicology and human toxicology*

Acute toxicity to fish	Good
Bioaccumulation	Good
Biodegradability	Good
Mutagenicity	Limited to good
Acute oral toxicity	Limited
Acute toxicity to algae	Limited
Acute toxicity to bacteria	Limited
Acute toxicity to mammals	Limited
Eye irritation	Limited
Hydrolisation	Limited
Photo-degradability	Limited
Skin irritation	Limited
Skin sensitisation	Limited
Carcinogenicity	Very limited
Chronic toxicity	Very limited
Teratogenicity	Very limited

*ecetoc: (Q)SARs: Evaluation of the commercially available software for human health and environmental endpoints with respect to chemical management applications. Technical Report No. 89

therefore seems possible that, within the next few years, we will be able to make predictions for some endpoints that are clearly defined in terms of mechanisms (such as acute toxicity, corrosive effects, irritant effect), and that they will be able to replace animal testing as part of an intelligent, tiered procedure[19].

Conclusion

It has become clear that neither the replacement methods nor methods based on molecular biology are advanced enough to make it possible to dispense completely with animal testing. On the other hand, it is impressive to see in concrete terms how selected examples illustrate just what progress has already been made in developing alternative methods to replace animal testing.

Annex IX of the Commission's REACH proposal explicitly provides for the possibility of data on toxicity to be acquired using replacement methods, on condition that these methods have been validated.

The German federal government explained in detail its position on the problem of animal testing under REACH back in spring 2004 in its response to Question 14 of the CDU/CSU parliamentary party's Major Interpellation – "The economic impact of the EU's chemicals policy". In this context it advocated:

- more, far-reaching regulation on data acquisition within the Community and on the use of alternative methods along the lines of the regulations already in place in Germany,
- the rapid validation of a number of other alternative methods that have already been developed and that could be used as part of the programme to phase-in existing chemicals, and
- encouraging systematic use of the data acquired in the initial phases for the further development of structure-activity analyses.

In line with these statements, the federal government has already fed into the Council's deliberations in Brussels on the Regulation extensive proposals for wording the text in a way that would close up regulatory gaps and consistently avoid multiple testing. With regard to the explanations of the data requirement annexes, it has also pushed for greater importance to be attached to alternative methods in the context of testing strategies for certain toxicological endpoints, for which this is possible.

In order to reduce the number of animal tests to be carried out, it will be decisive to what extent it proves possible in the next five years to bring to a successful and prompt conclusion the research efforts already begun on the development and validation of *in vitro* and *in silico* replacement methods.

However, we do not expect that complete replacement of animal testing will be achieved, particularly in the area of long-term effects of chemicals.

Correspondence to
Dr. rer. nat. Uwe Lahl
Director General
Federal Ministry for the Environment
Nature Conservation and Reactor Safety
Robert-Schumann-Platz 3
53175 Bonn, Germany
e-mail: Uwe.Lahl@bmu.bund.de

ICCVAM's Role in Validating *In Vitro* Test Methods for Endocrine Disruptor Screening*

Marilyn Wind[1], Bradley Blackard[2], Raymond Tice[2], William Stokes[3], David Hattan[4] and Leonard Schechtman[4]

[1]US Consumer Product Safety Commission, Bethesda, Maryland, USA; [2]Integrated Laboratory Systems, Inc./NICEATM, Research Triangle Park, North Carolina, USA; [3]National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM)/National Institute of Environmental Health Sciences Research, Research Triangle Park, North Carolina, USA; [4]US Food and Drug Administration, Rockville, Maryland, USA

Summary

Estrogen receptor (ER) and androgen receptor (AR) binding and transcriptional activation (TA) assays have been proposed as part of Tier 1 of the two-tiered endocrine disruptor screening battery the U.S. Environmental Protection Agency (EPA) is developing. ICCVAM comprehensively reviewed all in vitro ER and AR binding and TA assays and concluded that none were adequately validated. Minimum procedural standards such as dose selection criteria, number of replicates per test, appropriate positive and negative controls and criteria for an acceptable test were proposed that should be incorporated into standardised protocols for each of the four types of assays evaluated.

Keywords: endocrine disruptor screening, ER and AR binding tests, ER and AR transcriptional activation tests

Background

The Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) is an interagency group consisting of both U.S. Federal regulatory and non-regulatory agencies. Table 1 lists the 15 U.S. Federal agencies participating in ICCVAM.

ICCVAM was formed as an *ad hoc* committee in 1994 and was officially established by Congress in 1997. Its purpose is to: 1) Increase the efficiency and effectiveness of the U.S. Federal agency test method review; 2) Eliminate unnecessary duplicative efforts and share experience between U.S. Federal regulatory agencies; 3) Optimise utilisation of scientific expertise outside the U.S. Federal government; 4) Ensure that new and revised test methods are validated to meet the needs of U.S. Federal agencies and 5) Replace, reduce, or refine the use of animals in testing, where feasible.

Under the ICCVAM Authorization Act of 2000, ICCVAM's duties are defined as follows: 1) Review and evaluate new, revised or alternative test methods; 2) Facilitate interagency and international harmonisation of test methods; 3) Facilitate and provide guidance on test method development, validation criteria, and validation processes; 4) Facilitate acceptance of scientifically valid test methods; 5) Submit test recommendations to U.S. Federal agencies and 6) Consider petitions from the public for review and evaluation of validated test methods.

The National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM) is located at the National Institute of Environmental Health

*This paper reflects the views of the authors, has not been reviewed or approved by, and may not necessarily reflect the views of the Consumer Product Safety Commission or other agencies. This work was supported in part by NIEHS Contract N01-ES 35504. Since this paper was written as part of the official duties of the authors, it can be freely copied.

Tab. 1: ICCVAM participating U.S. Federal Agencies

Regulatory/Research	Non-Regulatory/Research
Consumer Product Safety Commission	Department of Defense
Department of Agriculture	Department of Energy
Department of Interior	National Cancer Institute
Department of Transportation	National Institute of Environmental Health Sciences
Environmental Protection Agency	National Institute for Occupational Safety and Health
Food and Drug Administration	National Library of Medicine
Occupational Safety and Health Administration	National Institutes of Health, OD
Agency for Toxic Substances and Disease Registry	

Sciences (NIEHS) in Research Triangle Park, NC. It functions as ICCVAM's operational arm, providing technical support. NICEATM conducts test method peer reviews and workshops and communicates and forms partnerships with stakeholders. All the background review documents referenced, the report of the peer review panel and documents on all the alternative test methods ICCVAM has reviewed can be found on our website (http://iccvam.niehs.nih.gov).

This paper focuses on the work ICCVAM has done on endocrine disruptor assays. In order to understand ICCVAM's role in validating *in vitro* test methods for endocrine disruptor screening, it is necessary to understand some of the history and why this is an area of interest. During the last two decades, there have been an increasing number of observations of reproductive and developmental abnormalities in animal populations exposed to high levels of certain persistent pollutants in the environment. There has also been an increased incidence of birth defects, cancers in hormonally receptive tissues and decreased fertility, which have been attributed to exposure of humans to endocrine disruptors. As a result of concern about the observations in both animals and humans, the U.S. Congress enacted The Food Quality Protection Act of 1996 and Amendments to the Safe Drinking Water Act in 1996. The U.S. Congress required the EPA to develop and validate a screening and testing programme to identify substances with endocrine disrupting (ED) activity. In 1998, EPA proposed their Endocrine Disruptor Screening Program (EDSP), which consists of a Tier 1 screening battery of *in vitro* and *in vivo* assays designed to identify substances capable of interacting with the endocrine system. Tier 2 of the EDSP is a battery of *in vivo* assays that provides detailed information on concentration response relationships and specific abnormal effects. Based on a weight-of-evidence evaluation of the results from the Tier 1 screening battery, Tier 2 *in vivo* tests are conducted. Included among the proposed Tier 1 *in vitro* assays are ER and AR binding and TA assays.

In April 2000, EPA requested ICCVAM to evaluate the validation status of *in vitro* ER and AR binding and TA assays. Background Review Documents were prepared for each type of assay. These documents can be found on the ICCVAM website. On May 21-22, 2002 (http://iccvam.niehs.nih.gov/docs/docs.htm#endocrine), an Expert Panel met in public session. In October 2002, a final report of the Expert Panel and the Endocrine Disrupter Working Group's (EDWG's) list of proposed substances for validation were made public for comment.

The Expert Panel was charged with the following:

- Review Background Review Documents and provide conclusions and recommendations on the following:
  - What assays should be considered for further evaluation in validation studies, and what is their relative priority
  - The adequacy of the proposed minimum procedural standards for each of the 4 types of assays
  - The adequacy of available test method protocols for assays recommended for validation studies
  - The adequacy and appropriateness of the substances/chemicals recommended for validation studies

Table 2 summarises what the databases for the ER and AR binding assays looked like.

ER binding was measured in fourteen different *in vitro* assays. These assays used ER derived from uterine cytosol from the mouse, rat and rabbit, from MCF-7 cells and MCF-7 cytosol and

Tab. 2: Estrogen and androgen receptor binding assays

	ER Binding Assay	AR Binding Assay
# of Assays	14	11
# Substances	635	108
% Substances Tested ≥ 2/Assay	37	31
# Chemical Classes Tested	17	5
Most Frequent Chemical Class	Polychlorinated Biphenyls	Nonphenolic Steroids
# Product Classes Tested	7	3
Most Frequent Product Class	Pharmaceuticals	Pharmaceuticals

Tab. 3: Transcriptional activation assays

	ER TA Assay	AR TA Assay
# of Assays	95	17
Total # of Substances Tested	703	146
# Tested for Agonism	634	109
% Agonists ≥ 2 Assays	36	45
# Tested for Antagonism	255	87
% Antagonists ≥ 2 Assays	37	26
# Chemical Classes Tested	15	7
Most Frequent Chemical Class	Polychlorinated Biphenyls	Nonphenolic Steroids
# Product Classes Tested	11	5
Most Frequent Product Class	Pesticides and Metabolites	Pharmaceuticals

from human cDNA clones of two human ER isoforms, ERα and ERβ (hERα and hERβ). Fusion proteins, in which glutathione (GST) was fused with the *def* domains of the human ERα (GST-hERα) and the ER from mice (GST-mER), chicken (GST-cER), anole (GST-aER) and rainbow trout (GST-rtER) were the basis for five assays. None of these assays had been validated and standard protocols were not used, even when the same assay was used in multiple laboratories. Although 635 chemicals had been tested in these assays, few were tested more than once in the same assay or in multiple assays. Only 8% of the substances were tested in seven or more assays. Thirty-seven percent of the substances were tested in two or more assays.

AR binding was measured in eleven different *in vitro* assays. These assays used AR derived from cytosol from rat prostate, rat epididymis and calf uteri, human cell lines (MCF-7, LnCaP) with endogenous AR, and a mammalian cell line (COS-1) transfected with human (h) AR. In addition, primary human genital fibroblasts (HGF) with endogenous AR, mammalian cell lines (COS-1) transfected with either hAR or rainbow trout ARα, and recombinant hAR Sf9 insect cells were also used in the assays. A majority (61%) of the 108 substances tested were only tested in one test. Since so few substances had been tested more than once in the same *in vitro* AR binding assay or in multiple assays using the same reference androgen, no quantitative or qualitative analyses of the comparative performance or the reliability of these assays was possible.

Table 3 summarises the database for the ER and AR TA assays. There are currently no generally accepted standardised methods for these assays. The *in vitro* TA assays used to identify ER agonists and antagonists fall into three broad groups: reporter gene assays using yeast cells; reporter gene assays using mammalian cells and cell proliferation assays using mammalian cells. Most of the mammalian cell lines and all the yeast cells lack an endogenous ER. The yeast strains and mammalian cell lines used in the various studies are listed in tables 2-1 to 2-3 of the Background Review Document for the ER TA assays. The ER used in the majority of the *in vitro* ER TA studies was human in origin. A few studies used ER derived from mouse or rainbow trout. A total of 703 substances had been tested in 95 *in vitro* ER TA assays. 634 substances were tested for agonism, with only 36% of the substances tested in two or more assays and only 8% tested in five or more assays. 255 substances were tested for antagonism; 37% percent of the substances were tested in two or more assays and only 3% were tested in five or more assays.

There are no standardised methods for performing AR TA assays. Studies were conducted using yeast (*Saccharomyces cerevisiae*), nine different mammalian cell lines, and one fish (carp) cell line. Of the mammalian cell lines used, six were human, two monkey and one was from Chinese hamster. The majority of studies used cells that were transiently transfected with AR. In other studies, cells were stably transfected with a plasmid containing the gene coding for the AR or contained an endogenous AR. The human AR was used in all but two of the studies. The remaining two studies used cells transfected with trout and mouse AR. One hundred forty-six substances were tested in 17 AR TA assays. Of the 109 substances tested for agonism, 45% were tested in two or more assays and 16% were tested in four or more assays. Of the 87 substances tested for antagonism, 26% were tested in two or more assays and 7% were tested in four or more assays.

Based upon a review of the background review documents, the expert panel concluded and ICCVAM concurred that:

1. There are no adequately validated *in vitro* ER- or AR-based assays.
2. No assays could serve as the basis for establishing performance standards.
3. There was little consistency among available protocols.
4. No test method protocol was adequately detailed and standardised.

In order to validate any of these test methods, the expert panel recommended minimum standards for all the assays; ICCVAM agreed with the recommendations. They are as follows:

- The limit concentration is 1 millimolar (mM); solubility characteristics must be taken into consideration.
- The concentration range should span at least 7 orders of magnitude and include at least 7 different concentrations.
- Triplicate measurements should be made at each concentration.
- For TA assays that use transient transfection methods, a constitutive reporter gene assay is needed to assess the efficiency of transfection.
- For TA assays, stability of cell lines with a stably transfected reporter should be monitored.
- For TA assays, cytotoxicity to define the upper limit for test substance concentrations is required.
- Reference estrogen/androgen and/or positive control responses must be consistent with historical data.
- For binding assays, substances that bind but do not bring about a 50% reduction in ER/AR binding should be classified as “equivocal”.
- For TA assays, a nonlinear regression model such as the Hill equation should be used to estimate the potency (EC_{50} or IC_{50} values) and slope of the concentration-response curve.
- Classification of a test substance as 'positive' should be based on statistical models pertinent to the characteristics of the assay.
- Replicate studies are not essential, but questionable data should be confirmed by re-testing.
- All studies requiring animals as tissue sources should be approved by an IACUC.
- The assays should be conducted following Good Laboratory Practice guidelines.

The following recommendations were made by ICCVAM for each of the four types of assays: For ER binding assays:

- Recombinant rat or human ER’s should be given highest priority for further test method standardisation, prevalidation and validation.
- An effort should be made to optimise a fluorescence-based method.
- Protocols should be standardised to incorporate minimum procedural standards.
- A minimum of 53 reference substances should be used in validation studies.

For AR binding assays, the following were recommended:

- Use recombinant protein as the source of AR.
- Standardise the protocol to incorporate the recommended minimum procedural standards.
- Conduct validation studies using, at a minimum, the 44 recommended reference substances.

For ER TA assays, the following were recommended:

- A comparative study to determine whether transiently or stably transfected lines are more appropriate should be conducted.
- Protocols should be standardised to incorporate minimum procedural standards.
- At minimum, assays should be validated using the designated 53 reference substances.

For AR TA assays the following were recommended:

- Develop a cell line containing an endogenous AR that is transduced with an adenovirus containing a reporter vector that shows high specificity for the AR.
- Standardise the assay protocol to include the recommended minimum procedural standards.
- Validate the assay using the same 44 reference substances recommended for the AR binding assays.

In addition to the recommendations for the specific types of assays, ICCVAM also recommended:

- No metabolic activation system should be recommended at this time.
- Prevalidation studies should be conducted to generate data necessary for biostatisticians to develop statistical methods for analysing data.
- The predictive value of these *in vitro* assays for estimating *in vivo* responses should be determined.
- A central repository of 78 substances should be organised.

Acknowledgments

All the expert panel members and especially those who chaired the individual panels (George Daston, John Stegman, Terry Brown and Elizabeth Wilson). The ICCVAM Endocrine Disruptor Working Group, especially the co-chair, David Hattan. The entire ICCVAM committee, especially Leonard Schechtman, the Chair, and William Stokes, the Executive Director. NICEATM staff without whose support we would not have had the background review documents, the expert panel meeting and all the documents on our website.

Correspondence to

Marilyn Wind, Ph.D.
Directorate for Health Sciences
US Consumer Product Safety Commission
4330 East West Highway
Bethesda, MD
USA 20814
e-mail: MWind@cpsc.gov

Workshop 5.11
Mechanisms of chemically-induced ocular injury and recovery

Ocular Toxicology *In Vitro* – Cell Based Assays

Monica Berry and Marcus Radburn-Smith
Academic Unit of Ophthalmology, University of Bristol, UK

Summary
Interactions between the three cell types in the cornea control differentiation and responses to stimuli. We have sequentially added cell types in a 3-dimensional construct to assess the minimal requirements for a toxicology model of the human cornea. Co-cultured cell types displayed patterns of cytokines different from the single cell-type 3D models. Following exposure to toxicants there were marked changes in cytokine profiles, related to the toxicant and markedly influenced by the epithelial cell-line used. For a rational choice of cell lines, their steady-state patterns of signal molecules should be compared to those in normal human preocular fluid.

Keywords: cornea, cytokines, human, in vitro, toxicology

Introduction and questions

The health of the ocular surface is intimately linked with quality of vision, and is also a determining factor in the success of surgical procedures. The ocular surface encompasses the outer structures of the eye (cornea, conjunctiva, and lids), the lacrimal gland and tear film and the innervation and immune structures that link them (Stern et al., 2004). Interactions within and between these components maintain the barrier to the environment and modulate responses to external stimuli (Cook et al., 2001; Stramer et al., 2003; Wilson et al., 2003; Hazlett, 2004; Holan et al., 2004; Wilson et al., 2004; Lema and Duran, 2005; Narayanan et al., 2005).

The Colipa Strategy for the Development of *in vitro* Alternative Methods focuses on physiological processes – from organ level to gene activation – that can be used to monitor responses to toxicants and recovery from injury. Among the programme aims, central to the replacement R of the 3Rs and to cell based assays, is identifying quantitative endpoints predictive of the nature and severity of injury. The ultimate aim is to replace the Draize test – currently the gold standard of ocular toxicity testing – with a scientifically valid animal-free strategy.

We needed to establish the minimal degree of complexity which is necessary and sufficient to represent the human cornea in cell culture models of toxicity. We have assessed the behaviour of undisturbed cultures and the response to a single toxicant application in constructs from monolayers to three-dimensional cultures comprising stratified human corneal epithelia and quiescent human stromal cells, to probe stability and correlates to injury and recovery.

Starting from the premise that depth (cell types) and area are the defining characteristics of an injury (Jester et al., 1998a; Jester et al., 1998b; Jester et al., 2000; Jester et al., 2001), we have explored endpoints related to metabolism and barrier function and sought correlations with patterns of cytokines, mediators that are known to be secreted by, and influence, the different cell types of the ocular surface. For the refinement R of the 3Rs and, more importantly, to anchor the *in vitro* tests with *in vivo* physiology, the add-on benefit of the latter endpoints is that signal molecules are readily detectable in (human) tears, and available through non-invasive tests.

Materials and methods

Epithelium

In addition to availability and stability with passage number, minimum *a priori* criteria were set for choosing human corneal

or conjunctival epithelial cell lines: ability to stratify and establish a barrier to penetration. Acute toxicity is correlated with barrier function (Konsoula and Barile, 2005). The corneal stroma is protected by the tight barrier of epithelium, which is in turn protected by the preocular fluid. A number of corneal and conjunctival epithelial cell lines were assessed for stratification, morphological differentiation and trans-epithelial electrical resistance (TEER).

Initially, epithelia were cultured according to originator's instructions: the cultures were then transferred and maintained in fully defined medium without antibiotics (KGM Bulletkit, Cambrex Biosciences, NJ, USA). Stratification was achieved by supplementing the medium with 1.25 mM $CaCl_2$ and cultured at liquid-air interface. The choice of serum-free medium follows from the need for minimally/non-activated stromal cells. This medium will not interfere with toxic effects and endpoint measurements.

Stroma

With approval of the Committee for Ethics in Research and consent for use in research, donor corneas unsuitable for transplantation were used to generate primary human corneal fibroblasts. These were amplified in DMEM with 10% foetal bovine serum (Invitrogen, Paisley, UK) and cryopreserved. Cells used to populate a collagen type I gel were maintained for at least 1 week in fully defined medium to inhibit cell proliferation and activated phenotype. Before seeding in collagen gels the cells were dissociated using trypsin. Soybean trypsin inhibitor (Sigma, Poole, UK) was used to stop the enzyme.

The matrix was built using 3 mg/ml bovine skin type I collagen (Vitrogen, Angiotech Biomaterials, Ca., USA), gelling at pH 7.4 (Taliana et al., 2000). Cells were included in the collagen solution at a concentration of 6 x 10^4 cells/ml. The gels, volume 0.25 ml per 24 well insert, were cast at their final placement and maintained in fully defined medium for 2 weeks before any further manipulation.

Epithelial-stromal construct

At least 2 weeks after matrix formation, epithelial cells were seeded at 10^5 cells/gel. The construct was maintained for 2 days and then lifted to liquid-air interface, where it was cultured for a further 7 days.

Toxicants and exposures

The choice of toxicants unifies the Colipa project: they are all from the same batch and are distributed to the collaborating laboratories by Colipa. We tested NaOH (Riedel-de-Haën, Germany), Sodium dodecyl sulphate (SLS, OmniPur®, Merck, Germany), and the non-ionic surfactant Tomadol 45-7 (Tomah, Ca. USA), each prepared in tissue culture water (Sigma). The concentration chosen was 0.66%, at which NaOH produces an initial lesion covering approximately three quarters of the surface of a stratified epithelium. For monolayers, this concentration was halved to take into account the smaller mass of the construct.

All constructs were exposed to 10 µl toxicant applied topically for 10 minutes and followed by an exhaustive wash with 1.0 ml tissue culture medium. Cultures were followed at 1 and 4 h for early signs of injury and then at 1, 3, 5, and 7 days post-exposure to evaluate the ensuing dynamics of responses to injury and recovery.

Assays

Trans-epithelial electrical resistance: TEER was measured with Millicel Electrodes (Millipore, Ma., USA) held in a fixed support. A cell-free insert was included as reference at each measurement.

Protein: After solubilisation with 1% Triton X-100 (Sigma) overnight at 4ºC, protein was quantified using the bicinchoninic acid kit (BCA™ Protein Assay Kit, Pierce, IL, USA) according to manufacturer's instructions.

Alamar Blue: This reagent (AlamarBlue, Serotec, Oxford, UK) was used at 2.5%, a concentration with minimal toxicity on repeat cell loading, every other day for 14 days. Fluorescence was measured using a Spectramax fluorimeter (Molecular Devices Corporation, Ca., USA) at $\lambda_{excitation}$ = 545 nm and $\lambda_{emission}$ = 590 nm, and expressed in arbitrary units. The sensitivity of the instrument was fixed for all readings.

Cytokines: TNF, IL-1β, IL-6, IL-8, IL-10 and IL-12p70 were measured simultaneously in each supernatant sample using the BD Cytometric Bead Array (BD Biosciences, Ca., USA) on a FACSCalibur (BD Biosciences). The fluorescence of samples and standard dilutions was assessed at least twice in each assay. Results are presented as means of a number of cultures, while the mean of replicate readings counted as 1. The dynamic range of the assays were 5-5000 pg/ml for IL-6 and IL-8.

Results

Some cell lines did not fulfil the minimal criteria for inclusion in tests and were not investigated further. For example, the immortalised human corneal epithelial cell line (Canadian cells, gift from May Griffith) reached only a fraction of the TEER achieved by the immortalised human corneal epithelial cells (USA cells, donated by Ilene Gipson), as shown in figure 1. The TEER of IOBA-normal human conjunctival cells (gift from Yolanda Diebold) was equal to that of the Canadian cells. Note that *in vivo* the conjunctiva is much leakier than the cornea.

It is interesting to note that the cytokine levels were specific to each cell line: in Araki-Sasaki supernatants IL-6 and IL-8 were in the order of 30 pg/ml and 400 pg/ml respectively, while in the Canadian cells levels were around 1500 and 2000 pg/ml respectively. In contrast, the USA cells secreted no detectable IL-6 when stratified. There was a clear time dependence of cytokine secretion, strongly influenced by the culture model. Stratification and culture at liquid-air interface, each affected the pattern and secretion of cytokines.

Further changes were observed when epithelia were cultured with keratocyte-populated stroma. An advantage of using a collagen matrix that can be altered by the cells growing therein becomes evident in that shrunken collagen gels can be easily identified. The difference in cytokine secretion from shrunken and unaltered gels is an indication of interactions between stro-

mal and epithelial cells (fig. 2), and stromal and endothelial cells (cell line gift from May Griffith). IL-8 concentrations also changed in response to the above factors. No cytokines were detected in supernatants of keratocyte gels.

When epithelial and endothelial cells were co-cultured with medium separating the two cell types, we observed a larger influence on IL-6 and IL-8 cytokine secretion than when in direct contact. Concentrations and magnitude of change with two cell types also depended on the medium used.

Preliminary experiments suggested that the optimum experimentation period would commence seven days after air-lifting and that the stratified epithelium would show no signs of attrition for seven days thereafter. We compared the different constructs for the projected duration of the experiment by measuring a metabolic correlate with Alamar Blue, protein as a measure of cell proliferation, and cytokine production. For monolayer cultures the starting point was taken at confluence, judged by microscopic inspection. While there was a continuous increase in cell mass (protein concentration) in the monolayer, the absolute redox activity remained stable, suggesting a decrease per cell. Protein concentrations and reductive activity were decreased in the USA stratified epithelia after day 5 of the experiment, while the epithelial-stromal constructs were stable for 7 days. Cytokine production per unit mass increased in monolayers, while in the stratified epithelia and the epithelial-stromal constructs cytokine levels were much lower and little changed during the experiment (fig. 3). Remarkably, IL-6 production was not quantifiable in stratified epithelia, and the levels of IL-8 were also below those of constructs containing a stromal cell seeded gel and stratified epithelia.

In response to single, short-term toxicant applications all constructs showed decreases in total reductive activity and cell mass (protein concentration). However, the time course and extent of these changes varied according to the construct: the non-ionic surfactant caused a very early and devastating decrease in cell number and activity in the monolayer. These losses of cells and metabolic capability occurred later in stratified epithelia; they were smaller and preceded by a short-term increase in reductive activity in epithelial-stroma constructs. In the case of this toxicant, an increasing secretion of IL-6 and IL-8 was measured towards the end of the experiment in the epithelial-stromal construct, perhaps suggestive of some recovery.

Discussion and conclusions

In response to public demand and European legislation, testing of cosmetic ingredients on live animals will cease in the near future. A consensus has emerged that any methodology that is developed to replace the use of animals in ocular toxicity testing has to fulfil a number of *a priori* criteria in order to be considered for pre-validation. These criteria include scientific purpose, mechanistic basis, statement of limitations and appropriate controls.

Within the Colipa Strategy for Development of *in vitro* Alternative Methods programme we have aimed at paradigms that address these criteria. Damage to the cornea is the most serious in functional visual terms, hence it was chosen as the target organ representing the ocular surface. We have reasoned that mediators of intercellular communication that can also be involved in inflammation would be a mechanistically-linked correlate of the response of the cornea to external stimuli. Inflammation is an aggravating result of ocular injury, and an integrating response of the ocular surface.

The results presented here indicate that cytokine production

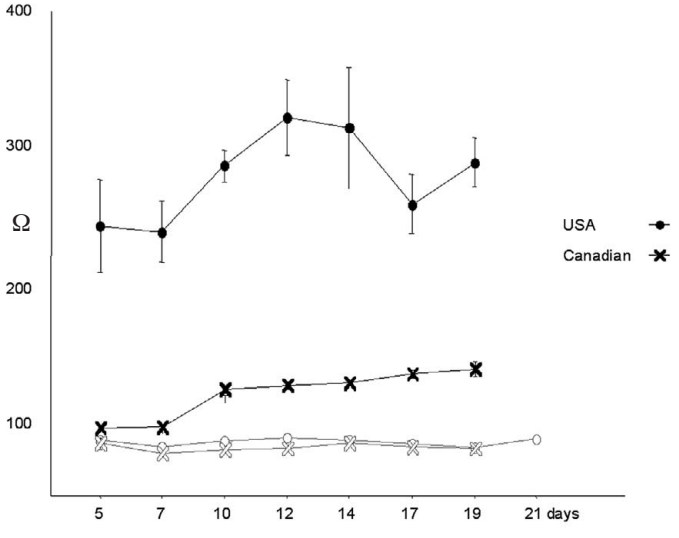


Fig. 1: Trans-epithelial electrical resistance
USA cells: filled circles; Canadian cells: filled crosses; blank inserts: grey, unfilled symbols

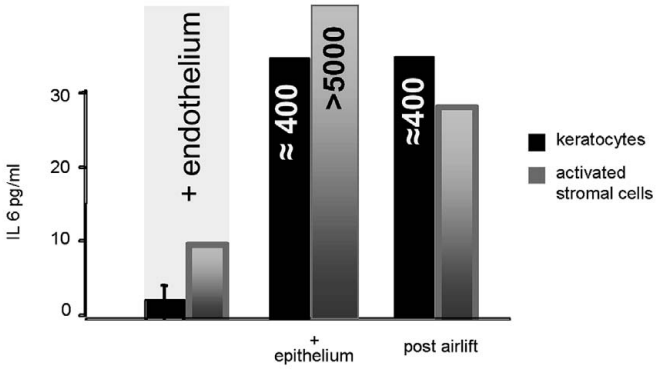


Fig. 2: IL-6 levels in different culture conditions
Black bars: quiescent keratocytes; grey bars: activated keratocytes (shrunken gel)

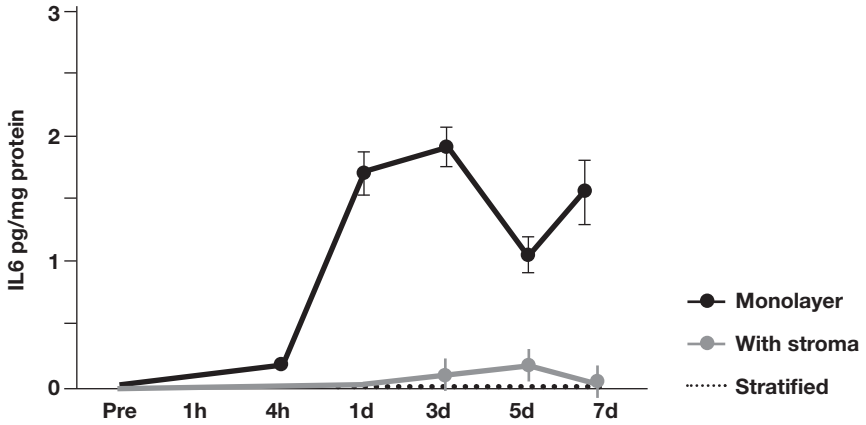


Fig. 3: IL-6 secretion in cell culture constructs with USA epithelia and quiescent keratocytes
Black line: monolayer; grey line: epithelialised stroma; dashed line: stratified epithelium

reflects the type and three-dimensional arrangement of cultured cells. Furthermore, the stability of the constructs is enhanced by the presence of different cell types. This indicates that a stratified epithelium alone is not a sufficient model for corneal toxicity tests. Elegant three-dimensional human corneal epithelial constructs, e.g. Skinethic (Van Goethem et al., 2005), could be combined with a three-dimensional stroma. Responses to a small number of chemicals suggest that the epithelial-stromal model allows some recovery to be studied and might be more discriminating than the single cell type models. In this construct, assessment of the time course of cytokine secretion after injury pointed to changes in cell activity that were not detectable by measuring reductive activity or cell mass.

There remains a caveat to the choice of cell lines in corneal constructs: each of the cell lines tested secreted different levels of cytokines. A comparison with data on cytokines in human tears is complicated by the fact that such data have not been obtained under standardised conditions, even when a healthy control group was included in the study. Our results indicate that after a chemical injury IL-6 increases in tears, compared to unexposed unrelated eyes (Berry and Jeffreys, 2001). In animals, IL-6 and IL-8 (and IL-1α) were elevated for a period of days after exposure to a toxicant, each cytokine taking a distinctive time course (Den et al., 2004). *In vitro*, twelve borderline irritation eye make-up removers increased IL-8 levels in supernatants of centrifuged CEPI monolayers from 1 to 23 times over control values (Debbasch et al., 2005), suggesting increased sensitivity of this endpoint to toxicant effects.

After sequentially building more complex corneal models and comparing their behaviour in steady state and after exposure to toxicants, we conclude that a three dimensional culture involving the three major cell types in the cornea is a stable construct and potential model for toxicity testing. Concentrations of toxicants used in these model-development studies have been at least one or two orders of magnitude smaller than those used in ex vivo studies done by our collaborators. This encourages us to further focus on mild to moderate toxicants when exploring the potential and limitations of this methodology.

References

Berry, M. and Jeffreys, D. (2001). Ocular injuries from household chemicals: Early signs as predictors of recovery. *In vitro and molecular toxicology 14*, 5-13.

Cook, E. B., Stahl, J. L., Lowe, L. et al. (2001). Simultaneous measurement of six cytokines in a single sample of human tears using microparticle-based flow cytometry: allergics vs. non-allergics. *J. Immunol. Methods 254*, 109-118.

Debbasch, C., Ebenhahn, C., Dami, N. et al. (2005). Eye irritation of low-irritant cosmetic formulations: correlation of in vitro results with clinical data and product composition. *Food Chem. Toxicol. 43*, 155-165.

Den, S., Sotozono, C., Kinoshita, S. and Ikeda, T. (2004). Efficacy of early systemic betamethasone or cyclosporin A after corneal alkali injury via inflammatory cytokine reduction. *Acta Ophthalmol. Scand. 82*, 195-199.

Hazlett, L. D. (2004). Corneal response to Pseudomonas aeruginosa infection. *Progress in Retinal and Eye Research 23*, 1-30.

Holan, V., Vitova, A., Pindjakova, J. et al. (2004). Corneal stromal cells selectively inhibit production of anti-inflammatory cytokines by activated T cells. *Clin. Exper. Immunol. 136*, 200-206.

Jester, J. V., Li, H. F., Petroll, W. M. H. et al. (1998a). Area and depth of surfactant-induced corneal injury correlates with cell death. *Invest. Ophthalmol. Vis. Sci. 39*, 922-936.

Jester, J. V., Li, L., Molai, A. and Maurer, J. K. (2001). Extent of initial corneal injury as a basis for alternative eye irritation tests. *Toxicol. In Vitro 15*, 115-130.

Jester, J. V., Molai, A., Petroll, W. M. et al. (2000). Quantitative characterization of acid- and alkali-induced corneal injury in the low-volume eye test. *Toxicol. Pathol. 28*, 668-678.

Jester, J. V., Petroll, W. M., Bean, J. et al. (1998b). Area and depth of surfactant-induced corneal injury predicts extent of subsequent ocular responses. *Invest. Ophthalmol. Vis. Sci. 39*, 2610-2625.

Konsoula, R. and Barile, F. A. (2005). Correlation of in vitro cytotoxicity with paracellular permeability in Caco-2 cells. *Toxicol. In Vitro 19*, 675-684.

Lema, I. and Duran, J. A. (2005). Inflammatory molecules in the tears of patients with keratoconus. *Ophthalmology 112*, 654-659.

Narayanan, S., Glasser, A., Hu, Y. S. and McDermott, A. M. (2005). The effect of interleukin-1 on cytokine gene expression by human corneal epithelial cells. *Exp. Eye Res. 80*, 175-183.

Stern, M. E., Gao, J., Siemasko, K. F. et al. (2004). The role of the lacrimal functional unit in the pathophysiology of dry eye. *Exp. Eye Res. 78*, 409-416.

Stramer, B. M., Zieske, J. D., Jung, J. C. et al. (2003). Molecular mechanisms controlling the fibrotic repair phenotype in cornea: Implications for surgical outcomes. *Invest. Ophthalmol. Vis. Sci. 44*, 4237-4246.

Taliana, L., Margaret, Evans, M. D. M., Slobodan, D. et al. (2000). Vitronectin or fibronectin is required for corneal fibroblast-seeded collagen gel contraction. *Invest. Ophthalmol. Vis. Sci. 41*, 103-109.

Van Goethem, Adriaens, E., Alepee, N. et al. (2005). Prevalidation of a new in vitro reconstituted human cornea model to assess the eye irritating potential of chemicals. *Toxicol. In Vitro*, in press.

Wilson, S. E., Mohan, R. R., Netto, M. et al. (2004). RANK, RANKL, OPG, and M-CSF expression in stromal cells during corneal wound healing. *Invest. Ophthalmol. Vis. Sci. 45*, 2201-2211.

Wilson, S. E., Netto, M. and Ambrosio, R. Jr. (2003). Corneal cells: Chatty in development, homeostasis, wound healing, and disease. *Am. J. of Ophthalmol. 136*, 530-536.

Acknowledgements

This project is part of the Colipa Strategy for the Development of *in vitro* Alternative Methods. We thank Colipa for financial support and the Eye Irritation Task Force members for the choice and procurement of toxicants and for many stimulating discussions throughout the project. We would like to thank Leon Bruner (Gillette), Richard Clothier (University of Nottingham), Yolanda Diebold and Margarita Calonge (IOBA, Valladolid), Dan Dimitrijevitch (University of North Texas), Ilene Gipson and Sandy Spurr-Michaud (Schepens Institute, Boston) and May Griffith (University of Ottawa) for gifts of cell lines and methodological advice. This project is part of the research undertaken by Marcus Radburn-Smith for a PhD degree at the University of Bristol. We are grateful to the National Eye Research Centre for additional support for this project.

Correspondence to

Dr Monica Berry
University of Bristol
Academic Unit of Ophthalmology
Bristol Eye Hospital
Lower Maudlin St
Bristol BS1 2LX
UK
e-mail: mon.berry@bristol.ac.uk

Can Toxicogenomics be Used to Identify Chemicals that Cause Ocular Injury?

Mike Boulton[1,2] and Mike Wride[2]

[1]Department of Ophthalmology & Visual Sciences, The University of Texas Medical Branch, Galveston, Texas, USA;
[2]School of Optometry and Vision Sciences, Cardiff University, Cardiff, UK

Summary

With the impending ban by the European Union on the use of the Draize eye test for cosmetic testing and to meet concerns regarding the 3Rs, an in vitro replacement is essential. This article considers a toxicogenomic approach that will allow a gene fingerprint to be produced from chemical exposure of corneal constructs prepared from human cell lines. Differential fingerprint profiles will allow identification of mild, intermediate and severe toxic preparations. This toxicogenomic approach will provide a rapid, high throughput, accurate diagnostic assay for the effects of toxic chemicals on the eye.

Keywords: Draize eye test, toxicogenomics, cornea

Introduction

The Draize eye irritation test has been mandated for routine use by manufacturers to study the safety of their products for marketing. This test, which is currently the gold standard, has several limitations:

a) It lacks objective and accurate quantification
b) It requires large numbers of animals
c) It is a slow, expensive, and a time consuming process
d) It is opposed by animal welfare groups as well as increasing numbers of the general public concerned about "non-medical" animal experimentation.

A total of 3,300 rabbits were used for eye irritation studies in the UK in 1994, of which the majority were used for cosmetics testing. It was recommended that the use of the Draize eye irritation test should be phased out within Europe by 1997, providing that suitable *in vitro* alternatives had been developed. Despite this not being the case, the European Parliament has subsequently decreed that the Draize eye test will no longer be allowed from 2009. This has left the manufacturing and cosmetic companies with a dilemma, since they must ensure safety of all new products entering the market. In an attempt to overcome these problems, alternative *in vitro* toxicity tests have been sought. Although a number of *in vitro* methods have been proposed as alternatives to the *in vivo* Draize test, they are either overcomplicated or oversimplified to such an extent that they measure only a single aspect of eye irritation (e.g. cell death) (Prinsen, 2005). Such tests are not considered sufficiently rigorous or informative, especially for moderate or mildly toxic chemicals. One possible alternative is to use a toxicogenomic approach to provide a gene fingerprint directory (i.e. pattern recognition), which will identify the toxicity of chemicals in corneal constructs. If successful, this will allow the production of specific diagnostic gene arrays for high throughput assays of all potentially toxic preparations that enter the market.

Chemical injuries to the eye

The first points of contact in chemical exposure to the eye are the eyelids, conjunctiva and cornea. Severe injury can result in corneal scarring and neovascularisation, which leads to opacification of the cornea, thus preventing light reaching the retina (Khaw et al., 2004). In extreme cases, treatment may require corneal transplantation, keratoprosthesis or limbal cell transplantation (Dogru and Tsubota, 2005). If these avenues are unsuccessful, the patient will suffer blindness or severe visual impairment. Mild chemical injury usually results in irritation and reddening. This either recovers by itself or requires some antibiotic and steroid treatment. Moderate chemicals initiate intermediate tissue damage and the outcome is dependent on the nature of the chemical (e.g. alkali, acid, detergent) and duration of exposure.

Not surprisingly, the nature of tissue damage and biological responses is very different between severe and mild chemical injury. Severe chemical injury to the cornea will result in tissue coagulation and large-scale cell death via necrosis. Immediate effects will be on the epithelium, but penetration into the stroma will cause damage to Bowman's membrane and the underlying stroma with loss of keratocytes and disorganisation of the regular collagen lamellae. This results in breakdown of barrier function, making the cornea more prone to infection, and destabilisation of corneal hydration leading to swelling. The release of cytokines and other mediators will attract macrophages, leading to an acute inflammatory response and neovascularisation. The immediate survival response is rapid proliferation of cells and disorganised matrix deposition, which often results in corneal scarring. Ulceration and corneal liquefaction can result, due to sustained upregulation of matrix metalloproteinases (MMPs). Not surprisingly, the biological responses associated with severe corneal damage result in a massive temporal change in gene expression. Furthermore, tissue remodelling can continue for up to 3 years post injury.

By contrast, mild chemical injury to the eye will result in limited tissue damage and a normal repair response. There is usually some cell loss, normally of the corneal epithelium, which is usually repaired within 72 hours through the division of resident transit amplifying cells and limbal stem cells. There will be some pain and irritation due to damage to the nerve endings in the epithelial layer. The mild associated inflammatory response will result in a reddening of the eye, which may need treatment with steroids and antibiotics. As with severe chemical injury, there will be a temporal change in gene expression, but the gene expression profile will be very different to that for severe injury.

While mild chemical injury does not normally impair vision, the discomfort to the individual is unacceptable and should be avoided if possible. Thus, it is important that manufacturers, employers and Health and Safety Executives are fully aware of potential chemical contraindications, can advise on protection and can recommend treatment following accidental exposure.

Feasible alternatives to the Draize test

Numerous alternatives to the Draize test have been proposed. These include cell culture, *ex vivo* organ culture and corneal reconstructs. While cell cultures (corneal or non-ocular, such as skin) can provide valuable information on severely toxic chemicals (the cells simply die!), they are limited in that they are usually composed of a single cell monolayer and, thus, do not measure tissue penetration. Such cultures have been used to assess the release of cytokines, growth factors and other mediators as a function of inflammation, but validation for mild or moderate chemicals is difficult due to the simplicity of the test.

Ex vivo organ culture using bovine, rabbit and porcine corneas has been attempted with outcome measures including morphology, release of biological mediators and opacification. Such models have the advantages that they have the typical 3D corneal structure, consist of numerous cell types and allow chemical penetration to be taken into account. Limitations include access to sufficient numbers of corneas, quality control (there will be considerable biological variation), the lack of a rapid throughput assay and the fact that these tissues are not human.

Corneal reconstructs (either epithelium/stroma or epithelium/stroma/endothelium) can be prepared in culture from established and well-characterised human cell lines (Griffith et al., 1999). These constructs, which produce an excellent 3D corneal architecture, allow excellent quality control and permit chemical penetration to be assessed. The limitation of these models to date is the nature of the outcome to be used to monitor mild or moderate chemical injury. Outcome measurements have tended to be similar to those described for organ culture and include morphological assessment and the measurement of the release of mediators. However, an alternative and more reliable outcome would be to monitor global gene expression changes in a construct following chemical exposure.

Aim

The strategy for the global gene expression approach is to develop a gene fingerprint directory (a pattern recognition approach), which will identify chemicals and preparations toxic to a bioengineered human cornea and that will be at least as sensitive as the Draize test.

Specific aims will include:

- Determination of different genetic fingerprints of human, bioengineered corneas exposed to different generic groups of toxic chemicals using microarray analysis.
- Recommendations for a set of selected human genes (<100) involved in eye irritation that are differentially expressed in the microarray experiments and are therefore important for the production of specific "diagnostic array" chips.

The proposed genomics assay for eye irritation will focus on pattern recognition rather than individual changes in genes. This pattern recognition approach, which depends on dynamics, dose and kinetics, will identify markers, e.g. for inflammatory processes, that may be exploited in the development of other *in vitro* assays.

Experimental design

For proof of principle, bioengineered human corneal constructs will be exposed to a range of generic chemicals (e.g. alkali, acid, detergent) with varying degrees of known ocular damage (severe, moderate, mild) and penetration, at varying concentrations, for different time periods. Affymetrix Gene Array chips will then be used to investigate changes in gene expression – using the U133A and B chips containing known and unknown genes (Wilson et al., 2002). A broad-based gene approach is essential in the first instance to avoid missing key diagnostic genes. A thorough statistical analysis using standard Affymetrix array analysis software will be undertaken and genes divided into clusters based on their functional categories. Statistically significant differences in gene expression profiles are expected for particular categories of genes under each experimental condition. Thus, it will be possible to define a gene expression fingerprint for the effect of each toxic chemical in this *in vitro* system to monitor tissue damage and repair.

Based on the gene expression profiles obtained, it is proposed to produce specific chips ("diagnostic arrays") containing selected genes that are differentially expressed in our human *in vitro* model microarray experiments. Thus, a rapid high-throughput diagnostic assay using bioengineered human corneas will be developed, which will give a global fingerprint of gene expression for the effects of toxic chemicals on the cornea. A subset of human genes will be selected that are consistently differentially expressed according to the arrays under the different experimental conditions. The possibility also exists to identify genes whose expression is consistently altered in different *in vitro* models (these genes will also be confirmed as differentially expressed

using RT-PCR and/or quantitative PCR). Individual genes and/or functional categories of genes, which are differentially expressed in all experimental paradigms, will be prioritised for inclusion on a human gene "diagnostic array", which will allow for identification of "toxicity fingerprints" in diagnostic arrays. Key diagnostic genes will include markers for inflammatory processes, apoptosis, and those for regeneration, and/or embryonic development. Selected genes will be arrayed onto glass slides for diagnostic array analysis. Appropriate array analysis software will then be used to compare the gene expression profiles obtained from diagnostic tests using cosmetics with the gene expression profiles generated by known toxic chemicals.
Of particular importance will be

1. the concentration of chemical to be tested, since potency is likely to vary significantly between *in vitro* and *in vivo*.
2. duration of exposure – minutes or hours?
3. number of exposures, should there be repeat exposures and should there be irrigation as would occur with tearing *in vivo*?
4. the timing of analysis. Since toxicity will involve both damage and repair, there will be a temporal change in the gene expression profile. Thus, the most comprehensive approach will be to look at expression profiles of early and late response genes.

Validation

The toxicogenomics model will require validation at a number of levels:

1. comparison with a database on chemical injury in humans
2. comparison with toxicity results obtained with the Draize test
3. approval by the relevant bodies (e.g. ECVAM, ICCVAM, NICEATM)

Conclusion

Toxicogenomics offers the possibility to produce a rapid, high throughput, accurate diagnostic assay for the effects of toxic chemicals on the eye. The approach, once developed and validated, will offer a high level of quality control, will be highly reliable, and will be quick and simple. Furthermore, this assay will meet the requirements of the European directive to phase out the Draize test by 2009 and should significantly reduce the number of animals used in this procedure.

References

Dogru, M. and Tsubota, K. (2005). Current concepts in ocular surface reconstruction. *Semin. Ophthalmol. 20 (2)*, 75-93.

Griffith, M., Osborne, R., Munger, R. et al. (1999). Functional human corneal equivalents constructed from cell lines. *Science 286 (5447)*, 2169-2172.

Khaw, P. T., Shah, P. and Elkington, A. R. (2004). Injury to the eye. *B.M.J 328 (7430)*, 36-38.

Prinsen, M. K. (2006). The Draize Eye Test and in vitro alternatives; a left-handed marriage? *Toxicol. In Vitro 20*, 78-81.

Wilson, A. S., Hobbs, B. G., Speed, T. P. and Rakoczy, P. E. (2002). The microarray: potential applications for ophthalmic research. *Mol. Vis. 8*, 259-270.

Acknowledgements

This study is funded by COLIPA.

Correspondence to

Professor Mike Boulton
Department of Ophthalmology & Visual Sciences
The University of Texas Medical Branch
301 University Blvd
Galveston
Texas 77555-1106
USA
e-mail: boultonm@utmb.edu

ICCVAM-NICEATM-ECVAM Symposium on Mechanisms of Chemically-Induced Ocular Injury and Recovery: Current Understanding and Knowledge Gaps

Karen Hamernik[1], Chantra Eskes[2], Jill Merrill[3], Neepa Choksi[4], David Allen[4], James Truax[4], Raymond Tice[5], Marilyn Wind[6], Leonard Schechtman[7] and William Stokes[5]
[1]US EPA*, [2]ECVAM, [3]US FDA, [4]ILS, [5]NICEATM/NIEHS, [6]US CPSC, [7]US FDA,

Summary
A scientific symposium on the Mechanisms of Chemically Induced Ocular Injury and Recovery was held on May 11-12, 2005 in the USA. It was organised and co-sponsored by ICCVAM, NICEATM/NIEHS, and ECVAM, with support from COLIPA. A major goal was to identify research needed to advance the development of test systems that meet regulatory testing requirements and provide for human health protection while reducing, refining (less pain and distress), and/or replacing the use of animals. This paper (one of three), provides a symposium overview and summarises the Session 3 discussions on mechanisms and biomarkers of chemically induced ocular injury and recovery.

Keywords: ocular injury and recovery, chemically-induced, mechanisms of ocular injury and recovery, chemically-related, ICCVAM-NICEATM-ECVAM

Presentation of symposium format and results

A brief overview of the entire symposium is presented in this paper. Also summarised is the part of the meeting (Session 3) dealing with discussions of issues and research needs related to the current understanding of mechanisms of ocular injury and recovery associated with chemical exposure. Other papers by Eskes et al. and Stokes et al., describe, respectively, the other two broad areas covered by the symposium, which were: *In Vitro* Models for Ocular Injury: Current and Potential Biomarkers, and *in Vivo* Models of Ocular Injury and Recovery: Current and Potential Biomarkers to Support Development and Validation of Predictive *in Vitro* Models.

Regulatory need to understand and assess ocular toxicity potential

Accidental eye injury is the leading cause of visual impairment in the United States, and workplace and household chemicals are a significant cause of these injuries, according to sources such as the American Academy of Ophthalmology and the US National Institute for Occupational Safety and Health. US Federal agencies charged with public health protection and the regulation of chemicals or pharmaceuticals and other types of products are concerned about being able to identify potential ocular hazards. Ocular safety and hazard testing in the United States had its origins in part stemming from a well documented case from the 1930s of an eyebrow and eyelash dye that caused severe effects including blindness in women (see illustration on the following US Food and Drug Administration website: http://www.fda.gov/oc/history/historyoffda/section2.html).

Symposium overview

The two-day May 2005 Symposium on the Mechanisms of Chemically Induced Ocular Injury and Recovery was organised and co-sponsored by the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM), the National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM), which is part of the US National Institute of Environmental Health Sciences (NIEHS), and the European Centre for the Validation of Alternative Methods (ECVAM). Additional support was provided by the European Cosmetic, Toiletry and Perfumery Association (COLIPA). The symposium was open to the public and was attended by 76 participants and speakers.

Representatives from ICCVAM, which is composed of 15 US Federal regulatory and/or research agencies, and members of the Ocular Toxicity Working Group (OTWG), one of the numerous ICCVAM subject matter working groups, assisted with symposium organisation, preparation and implementation. Liaisons from ECVAM to the OTWG also provided welcome assistance. A cadre of national and international invited experts from academia, industry, medicine, government, and the animal welfare advocacy area served as speakers and meeting co-chairs.

* The content of this paper does not represent the official position or policy of the US Environmental Protection Agency (US EPA) or any other US federal government agency

One overarching symposium aim was to review the state-of-the-science and current understanding of the pathophysiology, mechanisms and modes of action of chemically induced ocular injury, persistence and recovery. Another aim was to identify areas where research could be conducted to aid the development of test systems that would meet regulatory needs to provide for the protection of human health as well as be compatible with the promotion of animal welfare goals.

Specific symposium objectives were to:

1) Review current and potential molecular, cellular, tissue (e.g. histopathology), and clinical (e.g. corneal opacity, swelling, depth of injury, biomarkers of chemical injury and recovery and their usefulness for *in vivo* and *in vitro* testing models of ocular irritancy and corrosivity);

2) Identify knowledge gaps in the understanding of chemically induced ocular injury and recovery;

3) Identify and prioritise future research initiatives that would address current knowledge gaps and that are considered necessary to advance the development and validation of *in vitro* models of chemically induced ocular injury and recovery; and

4) Discuss and identify quantitative, objective endpoints that should be considered for inclusion in the current *in vivo* rabbit eye test and/or human clinical testing (e.g., more sensitive markers of injury and recovery) that would support development and validation of predictive *in vitro* methods and improve hazard characterisation and reliability.

The symposium agenda over the two-day period consisted of a total of five sessions. During Session 1, an overview of recent initiatives and US regulatory requirements for ocular toxicity testing was provided. Previous workshops on ocular injury and recovery were discussed, including: (1) the International Life Sciences Institute Health and Environmental Sciences Institute (ILSI/HESI) Workshop on Replacing the Draize Eye Irritation Test: Scientific Background and Research Needs, 1995, (Reference: *J. Toxicology – Cutaneous and Ocular Toxicology, 15(3)*, 211-234, 1996); (2) the ILSI/HESI Technical Committee on Alternatives to Animal Testing Expert Meeting on Eye Irritation Testing, Sept 29-30, 1996, (Reference: Ophthalmologic Perspectives on Eye Irritation Testing, Nussenblatt, et al., *J. Toxicology – Cutaneous and Ocular Toxicology 17*,103-109, 1998); (3) the COLIPA Workshop on Mechanisms of Eye Irritation, October 5-8, 1997 (Reference: Report on the COLIPA Workshop on Mechanisms of Eye Irritation, Bruner et al., *ATLA 26*, 811-820, 1998). Also mentioned was the ECVAM Workshop on Eye Irritation Testing: The Way Forward (Reference: Balls et al., *ATLA 27*, 53-77, 1998). Also noteworthy, are the results of two international workshops organised and sponsored by the [US] Interagency Regulatory Alternatives Group (IRAG) in 1991 and 1993, respectively: (1) Workshop on Updating Eye Irritation Test Methods: Proposals for Regulatory Consensus (Reference: *Food Chemical Toxicol. 31(2)*, 1993) and (2) Workshop on Eye Irritation Testing: Practical Applications of Non-Whole Animal Alternatives (Reference: *Food Chemical Toxicol. 35 (1)*, 1997).

Session 2 of the symposium reviewed current ocular injury and toxicity assessments and included a discussion of human chemically induced ocular injury. Session 3 (summarised in this paper) dealt with mechanisms and biomarkers of ocular injury and recovery. In Session 4, current *in vitro* models of ocular injury and recovery were discussed. The subject of Session 5 was *in vivo* quantitative objective endpoints to support development and validation of predictive *in vitro* models. Panel discussions were held after Sessions 3, 4 and 5. Each of the panels was given a set of pertinent questions to address based on the session topic, and the discussions were summarised at the end of the meeting.

Mechanisms and biomarkers of ocular injury and recovery

Symposium Session 3 focused on issues related to the current understanding of known mechanisms and modes of action of chemically related ocular injury, persistence and recovery. Speakers in this session discussed topics such as eye injury type and reversibility, cellular (e.g. corneal epithelium and endothelium) and tissue (e.g. cornea, conjunctiva, iris) responses to chemical injury in humans and animals, chemical-specific modes of action (e.g. for acids and bases) and the role of histopathology and depth of injury in evaluating ocular injury onset, extent, severity and recovery potential. Other speakers presented information on (1) the role of chemical toxicokinetics in ocular injury (including delayed onset of effect) and detoxification, (2) possible future roles for toxicogenomics in elucidating processes involved in ocular injury and its sequelae, and (3) the effects of chemical exposure on tear film.

Session 3 panel piscussion questions and responses

The panel for Session 3 was asked to respond to a series of four questions. Highlights of the panel's answers and opinions are presented below.

In the first question, the panel was asked to identify the currently known mechanisms and modes of action of chemically induced ocular injury and recovery. The panel responded that mechanisms of injury (e.g. cytotoxicity, protein coagulation, membrane saponification, disruption of the extracellular matrix, inflammatory cell infiltration with release of mediators, up-regulation of proteases and collagenases) are known for some chemicals and product types.

Further, the extent of ocular surface involvement and depth of corneal penetration may correlate with severity of lesions and recovery and thus could serve as a biomarker of reversibility potential. The panel proposed that more research could provide information on the general applicability of this concept. In addition, it was noted that some existing studies have shown species differences in response to the same chemical and dose-dependent differences in response to the same chemical (i.e. mechanisms of repair or injury at lower doses may not be relevant at higher doses). Currently, more data and information are available on severe ocular injury than for milder forms.

For the second question, the panel was asked to list current knowledge gaps in understanding of mechanisms and modes of

action of chemically induced ocular injuries and recovery. The panel responded that further assessment was needed of the relationship between type and severity of initial *in vivo* damage and persistence of effects. Furthermore, additional work was required to identify and develop for utility, biomarkers of injury and recovery (i.e. gene expression profiling, clustering and pathway analysis for ocular damage and repair) and to elucidate the role of tear film in ocular damage. It also would be important to improve the translation of *in vivo* physiology to *in vitro* models. This would involve enhancing current knowledge of such things as the role of metabolism and the linkage of expression of specific corneal proteins with injury and recovery processes.

In question three, the panel was asked to identify research initiatives needed to address current knowledge gaps and further characterise mechanisms and modes of action in order to advance the development and validation of predictive *in vitro* models of chemically induced ocular injury and recovery. The panel's opinion was that quantitative endpoints could be incorporated into the current *in vivo* test and evaluated to make the test more informative. These would include histopathology to correlate cellular changes with observational endpoints, HPLC and mass spectroscopy to evaluate penetration of substances into the eye, and depth of injury analysis to gain further insight on the utility of this measurement as a biomarker for reversibility/irreversibility of effects. The panel thought that further evaluation of species differences and dose-dependent differences in response was merited. Also mentioned was the need for additional work on tear film, especially with regard to composition determination, further elucidation of its role in protection of ocular constituents and the potential consequences of its disruption on the severity of effects of mild and moderate irritants. Other areas for additional work included the need to better characterise early onset versus delayed ocular responses to chemical agents, to investigate further the toxicokinetics of chemical exposure to the eye, to better elucidate the role of inflammatory responses in observed ocular damage and to better explore recovery mechanisms of the eye (such as effects on stem cells). Also suggested was additional evaluation of other *in vitro* models that might be more predictive or useful in hazard identification, such as human corneal models (isolated and reconstituted) and pig corneal models.

For the final question, the panel was asked to identify what *in vivo* biomarkers (e.g. molecular, cellular, morphological, clinical) should be further investigated as predictive indicators of severity of lesions, reversibility versus non-reversibility, or delayed responses. In answer, the panel mentioned as noteworthy: histopathology, quantitative endpoints obtained using standard biomicroscopy, confocal microscopy, selective staining, cytology, immunologic markers and gene expression profiling, and clustering and pathway analysis for ocular damage and repair.

A report of the Symposium proceedings will be available in the near future on the ICCVAM website (http://iccvam.niehs.nih.gov).

Correspondence to
Dr. Karen Hamernik (WHO Temporary Adviser)
Office of Science Coordination and Policy
US Environmental Protection Agency (EPA)
EPA East Building (Mail Code 7501M)
1200 Pennsylvania Avenue, NW
Washington, DC 20460
USA
e-mail: hamernik.karen@epamail.epa.gov

An Overview of the COLIPA Eye Irritation Research Programme

Penny Jones[1], Dan Bagley[2], Claudine Faller[3], Beatrice Le Varlet[4], Pauline McNamee[5], Irene Manou[6], Wolfgang Pape[7], Christine Van den Berghe[8] and Freddy Van Goethem[9]

[1]Unilever-Safety and Environmental Assurance Centre, Colworth Park, Sharnbrook, Bedford, UK; [2]Colgate-Palmolive, Piscataway, NJ, USA; [3]Cosmital Wella, Marly, Switzerland; [4]LVMH Recherche, Saint Jean de Braye, France; [5]The Procter & Gamble Company, Egham, Surrey, UK; [6]COLIPA, Brussels, Belgium; [7]Beiersdorf, Hamburg, Germany; [8]L'Oreal, Aulnay Sous Bois, France; [9]Johnson & Johnson, Pharmaceutical Research & Development, Beerse, Belgium

Summary

The COLIPA eye irritation programme incorporates research projects and collaborative activities with external partners. Three projects focus on understanding mechanisms of eye injury and identification of new in vitro endpoints that are more predictive of the in vivo human response to chemical injury: 1) investigation of whether kinetics/patterns of change in physiological function and signals of injury released from the cornea in vitro can predict a chemical's potential to damage the eye, with a focus on recovery; 2) identification of endpoints related to the magnitude of injury and quality of repair in human immortalised cells and 3D human conjunctival and corneal constructs; 3) a genomics project using a pattern recognition approach to identify new endpoints for injury/repair that builds on corneal models from projects 1 and 2. Collaboration of industry, academia, external scientific organisations and regulators is equally important. COLIPA is working with ECVAM by actively participating in its Eye Irritation Task Force and providing support for statistical analysis of current in vitro methods.

Keywords: in vitro, eye irritation, Draize eye test, alternative, replacement, COLIPA

Introduction

The efforts of the COLIPA (European Cosmetic, Toiletry and Perfumery Association) -SCAAT (Steering Committee on Alternatives to Animal Testing) Eye Irritation Task Force are dedicated to the development of *in vitro* alternative methods/strategies to assess eye irritation for the replacement of the Draize test. The research programme follows on from previous validation studies (e.g., Bagley et al., 1992; Balls et al., 1995; Brantom et al., 1997) and workshops (Bruner et al., 1998; Balls et al., 1999). The approach used incorporates integrated research projects and collaborative activities with external partners.

Research programme strategy and objectives

The approach of the COLIPA eye irritation research programme is to build on the experience of the earlier validation studies and scientific workshops through collaborations with academic institutions conducting COLIPA-supported projects and other collaborative activities with external partners. The objective is to gain an understanding of cellular molecular mechanisms of chemically induced eye irritation, with focus on corneal injury and recovery. Through this understanding the expected outcome is the identification of *in vitro* endpoints related to the dynamics of injury and recovery that are more predictive of the *in vivo* human response to chemical injury. This will enable the development of prediction models for pre-validation of new or improved *in vitro* methods that would proceed to formal validation.

There are three integrated research projects: 1) an investigation of whether kinetics/patterns of change in physiological function and signals of injury released from the cornea *in vitro* can predict a chemical's potential to damage the eye, with a focus on recovery; 2) identification of endpoints related to magnitude of injury and quality of repair in human immortalised cells and 3-dimensional human conjunctival and corneal constructs and 3) a genomics project using a pattern recognition approach to identify new endpoints for injury and repair that builds on the corneal models being evaluated in projects 1 and 2 for potential use in current/future *in vitro* assays. The approaches taken for each of these projects are given below.

Research Projects

Project 1: *In vitro* corneal culture eye irritation assay

This project is being conducted by Norbert Schrage and Markus Frenz at the University of Aachen, Germany, and was initiated in January 2002. The aims are 1) to develop an *in vitro* model of excised corneas maintained in culture to allow observation of injury and recovery following chemical exposure and 2) to investigate whether kinetics/patterns of change in physiological function and signals of injury released from the perfused cornea *in vitro* can predict a chemical's potential to damage the eye, with a focus on recovery.

The following stepwise approach has been adopted to develop a new isolated perfused corneal culture model that can be maintained for a period of time under steady culture conditions:

- Determination of viability and stability of the isolated perfused corneal culture system both morphologically and metabolically, and definition of the parameters to be used routinely to confirm system viability and stability
- Determination of suitability of the model to investigate wound healing by mechanical abrasion
- Exposure of the defined isolated perfused corneal system to model toxicants
- Identification of the morphological/biochemical markers of injury/recovery to be used routinely to evaluate toxicant effects
- Investigation of evaluation methods including biomicroscopy, pachymetry and glucose/lactate turnover for system viability/stability, LDH, cytokines (IL-1α, IL-2, IL-6 IL-8, MIP1) and growth factors (FGF, VEGF) for evaluation of dynamics of injury and recovery after mechanical trauma or toxicant exposure.

Project 2: Cell culture models for ocular toxicity studies

This project is being undertaken by Monica Berry and Marcus Radburn-Smith at the University of Bristol, UK, and was initiated in January 2002. The aims of this project are 1) to sequentially build 3D human corneal constructs consisting of epithelium, stroma and endothelium in order to better understand underlying mechanisms of action of eye irritation and 2) to identify new endpoints related to magnitude of injury and quality of repair in human corneal models that will enable prediction of the nature and severity of toxicant effects.

This following approach is therefore being used in this project to investigate the physiological responses to ocular injury (e.g. cell activation/signalling to immune system effector cells) by evaluating responses to model toxicants in increasingly complex corneal constructs:

- Investigation of human corneal and conjunctival cell lines for culture conditions, growth characteristics and suitability for use in 3D constructs
- Development of stratified epithelia, stromal equivalents and construction of two layer models (epithelium and stroma)
- Construction of three layer models by the addition of an endothelium cell layer
- Exposure of monolayers and stratified models to model toxicants using the following evaluation methods
- Light and confocal microscopy, characterisation of surface markers and differentiation, barrier formation assessment, membrane damage, metabolic activity, profiling of cytokine secretion.

Project 3: Development of gene expression fingerprints to identify toxic damage to the cornea

This project will be initiated in the summer of 2005 and will be led by Mike Boulton, University of Cardiff, U.K. The aims of this project are 1) to generate proof of concept that generic chemicals will cause differential gene expression in human bioengineered corneas 2) to identify gene expression profiles in bioengineered human corneas exposed to generic classes of chemicals and 3) to develop a gene fingerprint directory to identify chemicals toxic to a bioengineered human cornea. The principal outcome should be the application of the knowledge to better understand new endpoints for eye irritation and enable further development of current and future *in vitro* methods.

This project will use the following approach:

- Microarray analysis with human Affymetrix Gene Array chips using RNA from untreated versus treated (model toxicants) bioengineered human corneal constructs developed in the Bristol project
- Analysis expansion to define a gene expression fingerprint for the effects of each chemical class. Fingerprints will be generated for two time points to assess both injury and repair.

External collaboration

Equally important to achieve validated *in vitro* methods is collaboration between industry, academia, external scientific organisations and regulators. COLIPA is working with ECVAM by active mutual participation in both COLIPA and ECVAM Eye Irritation Task Forces to ensure that the research efforts are synergistic rather than duplicative between the two organisations. The Task Force also uses external contacts in order to benchmark and refine the eye irritation research projects.

References

Bagley, D. M., Bruner, L. H., de Silva, O. et al. (1992). An evaluation of five potential alternatives in vitro to the rabbit eye irritation test in vivo. *Toxicology in Vitro 6*, 275-284.

Balls, M,, Botham, P. A., Bruner, L. H. and Spielmann, H. (1995). The EC/HO international validation study on alternatives to the Draize eye irritation test. *Toxicology In Vitro 9*, 871-929.

Balls, M., Berg, N., Bruner, L. H. et al. (1999). The report and recommendations of ECVAM workshop 34. *ATLA 27*, 53-77.

Brantom, P. G., Bruner, L. H., Chamberlain, M. et al. (1997). A summary report of the COLIPA international validation study on alternatives to the Draize rabbit eye irritation test. *Toxicology in Vitro 11*, 141-179.

Bruner, L. H., de Silva, O., Earl, L. K. et al. (1998). Report on the COLIPA workshop on mechanisms of eye irritation. *ATLA 26*, 811-820.

Acknowledgments

The Eye Irritation Taskforce members would like to acknowledge the support of COLIPA-SCAAT for this research programme and the collaborating researchers at the Universities of Aachen, Bristol and Cardiff.

Correspondence to

Penny Jones
Safety and Environmental Assurance Centre
Colworth Park
Sharnbrook, Bedford UK
e-mail: Penny.Jones@unilever.com

Session 5.02
New approaches to risk assessments (ESTIV-Session)

In Vitro-In Vivo Extrapolation of Toxic Potencies for Hazard and Risk Assessment – Problems and New Developments

Michael Gülden and Hasso Seibert
University Medical School Schleswig-Holstein, Kiel, Germany

Summary
The aim of toxicological hazard assessment is to characterise the dangerous properties of chemicals for man and the environment. Information on both (a) the toxic potential, i.e. the spectrum of toxic effects a chemical can produce, and (b) the toxic potency, i.e. the quantitative relationship between dose/concentration and toxicity, are essential to characterise the toxic hazard. Toxicological risk assessment comprises hazard assessment and aims to characterise the likelihood and severity of adverse effects occurring to man or the environment following exposure to a chemical under defined conditions.
Two fundamental problems hamper the application of in vitro assays for hazard assessment: firstly, the endpoints of toxic action detectable in vitro are less complex and, importantly, mostly different from those assessed in vivo (toxicodynamic problem). Secondly, toxic concentrations determined in vitro are not equivalent to toxic doses or concentrations in vivo. This is due to important differences in biokinetics and bioavailability of chemicals in vitro and in vivo (toxicokinetic problem).
This contribution is focussed on the second aspect. It is demonstrated, how predictions can be made of the toxic concentrations in human serum or the aquatic environment, which are equivalent to toxic concentrations in vitro. This can be achieved by the application of a recently developed quantitative extrapolation model, taking into account substance and system specific parameters important for the bioavailability of chemicals. It appears that this approach represents a real progress in solving part of the "toxicokinetic problem".

Keywords: bioavailability in vitro, free concentrations, in vitro – in vivo prediction model, equivalent exposure approach, acute toxicity

Introduction

Toxicological hazard and risk assessment until today largely rely on animal toxicity data. Animal toxicity tests are performed to identify the toxic effects that a chemical can induce (hazard identification) and to determine the toxic potency, i.e. the quantitative relationship between external exposure level (administered dose, environmental concentration) and the incidence and severity of toxic effects (dose-response assessment). Both components of hazard assessment are essential for toxic hazard classification and labelling, determination of safe exposure levels and, in connection with exposure assessment, for risk assessment (fig. 1). The aim of toxicological risk assessment is to characterise the likelihood, the nature and the severity of adverse effects occurring to man or the environment caused by defined exposure to a chemical. If *in vitro* methods shall be used for toxicological hazard and risk assessment and for safety evaluation, they must provide information which can be used to reliably assess the kinds of toxic hazards and the corresponding toxic potencies.

Despite considerable efforts during the last about 20 years, *in vitro* methods have only been accepted for hazard identification and classification in a few cases, e.g. in testing for genotoxicity/mutagenicity, acute local skin and eye toxicity and phototox-

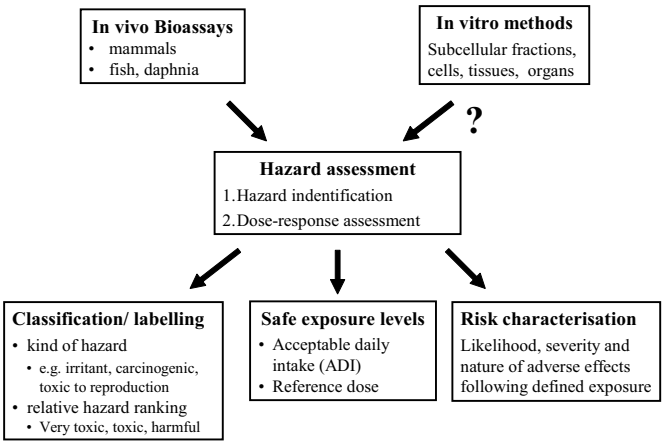


Fig. 1: *In vitro* methods in hazard and risk assessment.

icity. Currently, no *in vitro* methods are available that can be used to identify more complex toxic hazards, like carcinogenicity, reproductive toxicity, endocrine disruption, immunotoxicity or acute and chronic systemic toxicity.

The two fundamental problems which hamper the application of *in vitro* toxicity data for hazard assessment are well known. The first problem, which may be called the "toxicodynamic problem", is that the types of toxicity assessed *in vitro* and *in vivo* are not the same. Endpoints of toxic action detectable *in vitro* are much less complex and, importantly, mostly different from those assessed *in vivo*. No correlates for responses of organisms like death or disease, reproductive and developmental failure, cancer development or impairment of nervous, endocrine or immune system functions can be assessed *in vitro*. The diversity of mechanisms and interactions between tissues and organs involved in these complex toxic endpoints cannot be modelled by simple *in vitro* systems. *In vitro*, interactions or responses at cellular or subcellular level can be measured which are at best mechanistically connected to the observed *in vivo* phenomena. It is widely assumed that this toxicodynamic problem may be solved, at least in part, by the use of *in vitro* test batteries and tiered testing schemes, covering diverse mechanisms of action (e.g. Seibert et al., 1994; Seibert et al., 1996; Worth and Balls, 2002). The question here is how to identify relevant in vitro tests and how to combine their data.

The second problem, which may be called the "toxicokinetic problem", is that the exposure levels that characterise the potencies *in vitro* and *in vivo* are not equivalent. Exposure levels *in vivo* are given, for instance, as concentrations in the ambient air or water, or as administered doses, while exposure levels *in vitro* are most often given as nominal concentrations in the medium surrounding the cells or subcellular components. *In vivo* the effectivity of exposure and the bioavailability of chemicals is largely determined by biokinetic processes, which are either absent or quite different *in vitro*. That is why *in vitro* and *in vivo* exposure levels are not equivalent. If *in vitro* toxic potency data shall be used for *in vivo* toxic potency assessment, they have to be transformed into the external doses or concentrations *in vivo*, which are equivalent to the toxic exposure levels *in vitro*.

A further complication is that the nominal effective concentration used to characterise the *in vitro* toxic potency of a chemical for a given cell type and endpoint is not necessarily representative for the activity of that chemical but may vary with the composition of the *in vitro* system used, for instance with the albumin or serum concentration in the culture medium or the cell density (e.g. Bohets et al., 1994; Finlay and Baguley, 2000; Hestermann, 2000; Gülden et al., 2001; Seibert et al., 2002; Heringa et al., 2004). Thus, the question is: which *in vitro* potency measurement should be used for extrapolation to *in vivo* potency?

This contribution is focussed on some aspects of the "toxicokinetic problem". Firstly, a recently developed method (equilibrium distribution modelling) is described which is used to obtain composition independent, representative *in vitro* potency measurements. Secondly, a strategy is outlined for extrapolation of the internal or external exposure levels, which are equivalent to *in vitro* toxic concentrations (equivalent exposure approach). Finally, two examples for the application of this approach are presented to demonstrate its practical value.

Distribution model for chemicals in cell cultures

Cell cultures consist of at least three components: culture vessel, cells, and medium, the latter of which is frequently supplemented with 5-20% serum. It is known that chemicals can be bound to proteins and accumulated in lipids and that this occurs to a great extent in serum. Additionally, cells can concentrate chemicals via diverse mechanisms. Using "negligible depletion" solid-phase microextraction (nd-SPME) it was shown that *in vitro* considerable amounts of added chemicals can be bound by serum or cells and thus become unavailable for toxic action (Vaes et al., 1997; Heringa et al., 2004). Only the aqueous concentration of the freely dissolved chemical is considered to be related to the intensity of a pharmacological or toxic effect. The above-mentioned dependence of *in vitro* toxic potencies on albumin, serum or cell concentration in the *in vitro* systems is easily explained by an altered availability of the nominal concentrations of the chemicals when the concentration of the binding components is changed. A distribution model for chemicals in cell cultures has been developed to gain a composition independent measure of *in vitro* toxic potency, i.e. the free effective concentration (Gülden and Seibert, 1997; Gülden et al., 2001; Seibert et al., 2002).

This model is based on the assumptions that the total substance added is equilibrated between cells and extracellular components, the equilibrium concentrations are constant throughout the exposure period, and the total cell volume is negligible compared to the volume of the extracellular medium. Furthermore, it is assumed that chemicals can be freely dissolved in the extracellular and cellular water, can be bound to extracellular proteins and can partition into cellular and extracellular lipids. Any substantial influence of cell binding other than partitioning into cellular lipids is neglected. Cellular and extracellular lipids are treated as a single compartment.

Then, the nominal effective concentration (EC_{50}) can be expressed as a function of the free effective concentration

(ECu_{50}), the protein bound concentration (Cb), the concentration in the lipid compartment (C_L) and the relative volume of lipid (V'_L) in the *in vitro* system:

$$EC_{50} = ECu_{50} + Cb + C_L \bullet V'_L \quad (1)$$

Applying the Nernst distribution law to describe partitioning into lipids using the octanol/water partition coefficient (K_{ow}) and the mass action law to describe binding to extracellular proteins represented by albumin, the most abundant and most important binding protein in serum, the following equation was developed:

$$ECu_{50} = \frac{EC_{50} - B \bullet P}{1 + K_{ow} \bullet V'_L} \quad (2)$$

Provided that the albumin concentration in the medium (P), the relative lipid volume in the *in vitro* system, the specific binding of the chemical to albumin (B) and its octanol/water partition coefficient are known, this equation enables the calculation of the free effective concentration from the nominal effective concentration. The specific binding to albumin can be determined from EC_{50}-measurements at different albumin concentrations in the medium (Gülden et al., 2003).

The distribution model has been applied to characterise the distribution and availability of cytotoxic concentrations of a variety of chemicals in the mouse Balb/c 3T3 cell cytotoxicity test system (Gülden et al., 2002, Gülden and Seibert, 2005). Despite the low albumin concentration of about 18 µM in culture medium supplemented with 5% foetal bovine serum and the low lipid content ($V'_L = 10^{-4}$ l/l), various compounds have been found, whose availability is considerably, in some cases drastically, lowered by albumin binding and/or partitioning into lipids (tab. 1).

In vitro-in vivo extrapolation of toxic potencies: The equivalent exposure concept

In order to use *in vitro* toxic potency data for *in vivo* toxic potency assessment, the differences in biokinetics *in vitro* and *in vivo* have to be taken into account. *In vivo* biokinetic processes, like resorption, biotransformation, excretion and distribution, determine the relation between the administered toxic dose or concentration of a chemical in the ambient air or water and the internal dose or the actually effective free concentration at the target site (fig. 2). As outlined above, the free effective and not the nominal effective concentrations must be considered to be the representative, composition independent *in vitro* potency measurement. From a toxicodynamic point of view, nominal concentrations and doses of a chemical can be regarded to be equivalent if they are connected to the same free concentration of that chemical. Thus, it can be taken as a basic demand on an *in vitro-in vivo* extrapolation model for toxic potencies that it is aimed at determining external or internal doses/concentrations of a chemical *in vivo* that result in a free concentration which is comparable to the *in vitro* free effective concentration. We call this the equivalent exposure concept.

Tab. 1: Nominal (EC_{50}) and free (ECu_{50}) cytotoxic concentrations of chemicals and their free fractions (fu) in the Balb/c 3T3 cytotoxicity test system (Gülden et al., 2002; 2003).

Substance	EC_{50} (µM)[a]	ECu_{50} (µM)[b]	fu[c]
Pesticides			
p,p´-DDE	34.4	0.034	0.001
p,p´-DDT	30.6 ± 2.4	0.035	0.0011
Dieldrin	33.3 ± 7.3	9.27	0.28
Lindane	65.3	35.5	0.54
Malathion	60.7	53.7	0.88
Paraquat	74.6	74.6	1.0
2,4-Dichlorophenoxy acetic acid	877	803	0.92
Phenols			
4-Octylphenol	18.6 ± 3.6	2.73	0.15
4-Nonylphenol	24.1 ± 6.9	3.36	0.14
Hexachlorophene	4.06	0.027	0.0067
Pentachlorophenol	39.2 ± 13	≤ 2	≤ 0.050
2,4,5-Trichlorophenol	64.4	17.7	0.27
2,4-Dichlorophenol	281	235	0.84
4-Chlorophenol	219 ± 44	192	0.88
Phenol	1140	1140	1.0

a : Data are means of at least two independent measurements ± SD if n ≥ 3.
b: Calculated with Eq. (6) using a relative lipid volume $V'_L=10^{-4}$ l/l, an albumin concentration P = 18 µM, the pH 7.4 adjusted values for K_{ow} obtained from the literature and the values for B, the specific binding to albumin, as determined from EC_{50}-measurements in the presence of different albumin concentrations.
c : $fu = ECu_{50}/EC_{50}$

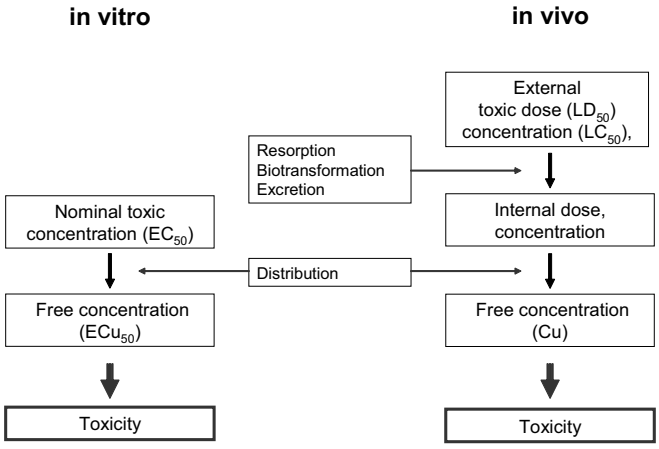


Fig. 2: Influence of biokinetic processes on the bioavailability of chemicals *in vitro* and *in vivo*.

In the following, using cytotoxic potency data obtained with the mouse Balb/c 3T3 cytotoxicity test system, the application of this approach to predict (a) fish acute toxic potencies and (b) toxic human serum concentrations is demonstrated.

Application example 1: Prediction of acute aquatic toxicity to fish

Fish acute toxicity tests are performed in large numbers for toxicological hazard and risk assessment of chemicals in the aquatic environment. Some efforts have been undertaken to develop *in vitro* cytotoxicity tests with fish cell lines as alternatives to the acute fish test (for review see Segner, 1998; Castano et al., 2003). Using cell death as an endpoint of cytotoxicity, various fish cell lines have been used to measure the cytotoxic potency of a variety of chemical compounds. In any case, fish cell line tests showed a clearly lower sensitivity than the *in vivo* fish tests when *in vitro* cytotoxic concentrations (EC_{50}-values) of groups of chemicals were compared with their acute toxic concentrations (LC_{50} values) (Castano et al., 2003; Segner, 2004). The low sensitivity of fish cell line tests hampers their use as alternatives to acute fish tests.

An example using cytotoxicity data obtained with the rainbow trout R1 cell line (data from Segner and Lenz, 1993; Segner, 2004) is shown in figure 3. It can be seen that only for a few substances the *in vitro* cytotoxic concentrations are close to the acute toxic concentrations or somewhat lower. In most cases the cytotoxic concentrations are higher, up to more than a factor of 100. The deviations between LC_{50}- and EC_{50}-values become more frequent and more pronounced with increasing potency.

Meanwhile, various studies have shown that this "low sensitivity" is not a unique feature of fish cell lines (Segner 2004, Castano and Gómez-Lechon, 2005, Gülden et al., 2005). In fact, the sensitivity of mammalian and human cell lines towards the cytotoxic action of chemicals, in general, is comparable to that of fish cell lines, provided that the experimental conditions are similar. The sensitivity of *in vitro* cytotoxicity tests using cell lines, however, can be increased if cell growth inhibition instead of cell death is used as endpoint for cytotoxicity (Gülden et al., 2005). Nevertheless, the principal relation between *in vitro* and *in vivo* assays, i.e., the lower sensitivity of *in vitro* assays especially for chemicals with higher toxicity, remains.

We have investigated whether a reduced availability of chemicals *in vitro* can account for the remaining lower sensitivity of *in vitro* cytotoxicity test systems (Gülden and Seibert, 2005). Fish cell lines are usually cultured in the same basal culture media as mammalian cell lines, also supplemented with 5-10% bovine serum. It has been shown above that, under these conditions, considerable fractions of nominal effective concentrations of chemicals can be bound to extracellular serum albumin and can be accumulated in cellular and extracellular lipids. The bioavailable effective concentrations of chemicals (ECu_{50}), thus, can be much lower than their nominal effective concentrations (EC_{50}). On the other hand, the fish acute toxic concentrations (LC_{50}) ideally refer to the bioavailable aqueous concentration of the freely dissolved chemicals. This is warranted by quantitative analytical chemistry and/or proper design of the fish acute toxicity test if performed according to the OECD test guidelines. This suggests that the free and not the nominal effective concentrations *in vitro* are the appropriate *in vitro* potencies for aquatic toxicity assessment.

To test this hypothesis, nominal cytotoxic concentrations of chemicals determined with the Balb/c 3T3 proliferation inhibition assay were converted into free effective concentrations using the *in vitro* distribution model outlined above (Eq. 2). The octanol/water partition coefficients were taken from the literature and values for the specific binding to albumin were determined from EC_{50}-measurements in the presence of different albumin concentrations with the Balb/c 3T3 test system. Organic chemicals covering a wide range of cytotoxicity and lipophilicity were selected for which acute fish toxicity data were avail-

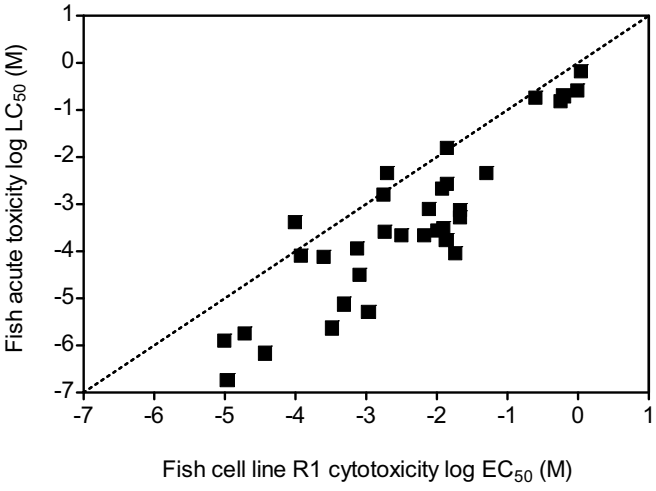


Fig. 3: Comparison of the aquatic acute fish toxic concentrations of chemicals with their *in vitro* cytotoxic concentrations in the R1 fish-cell line. Data from Segner and Lenz (1993) and Segner (2004).

able in the Ecotoxicology Database (ECOTOX) of the US EPA for at least one of the three fish species, medaka (*Oryzias latipes*), rainbow trout (*Oncorhynchus mykiss*) and fathead minnow (*Pimephales promelas*).

The results show that the sensitivity of the Balb/c 3T3 cytotoxicity assay and the correspondence between *in vivo* and *in vitro* toxic potencies can clearly be increased when the free cytotoxic concentrations instead of the nominal cytotoxic concentrations are used as measure of cytotoxic potency. An example is shown in figure 4 with data taken from Gülden and Seibert (2005). The few remaining prominent underestimations of acute fish toxic potencies can be explained by more specific mechanisms of acute toxic action than basal cytotoxicity. Chemicals, like the tested nicotine or the pesticides dieldrin and lindane are known to exert acute toxic action via specific interference with neuronal transmission. If a chemical exerts acute toxicity via interference with specific cell functions at concentrations lower than those that are cytotoxic, its acute toxic potency must be higher than its cytotoxic potency. That means, even if the toxicokinetic problem is solved, cytotoxicity testing alone cannot be expected to safely assess the acute aquatic toxicity to fish. However, since cytotoxicity can be considered with certainty to be a cause for acute toxicity, cytotoxicity testing *in vitro* may well be suited to assess the minimal acute fish toxic potency of chemicals.

Practical *in vitro* test strategy for estimating minimal aquatic toxic potency

Based on these results and earlier investigations on the impact of protein binding and partitioning into lipids on the availability and toxic potency of chemicals *in vitro* (Gülden et al., 2001 and 2003; Seibert et al, 2002), the following practical *in vitro* testing strategy for estimating the minimal aquatic toxic potency of chemicals is proposed.

1) Determination of nominal EC_{50}-values for growth inhibition using a mammalian cell line.

2) Determination of the distribution relevant composition of the *in vitro* test system used, like albumin concentration in the medium (P) and the relative volume of cellular plus extracellular lipids (V'_L).

3) Determination of the octanol/water-partition coefficient (K_{ow}) of the tested chemicals. If the octanol/water partition coefficient of the chemical is lower than the reciprocal of the relative lipid volume in the *in vitro* system ($Kow < 1/V'_L$), partitioning into lipids is not likely to affect the availability of the chemical in the *in vitro* system.

4) If the EC_{50}-value of a chemical is much higher than the albumin concentration in the *in vitro* system ($EC_{50} \gg P$), quantitatively relevant binding to albumin is not likely to occur. Then, the free effective cytotoxic concentration *in vitro* is given by the following equations:

a) If $K_{ow} < 1/V'_L$

$$ECu_{50} = EC_{50} \quad (3)$$

b) If $K_{ow} \geq 1/V'_L$

$$ECu_{50} = \frac{EC_{50}}{1 + K_{ow} \bullet V'_L} \quad (4)$$

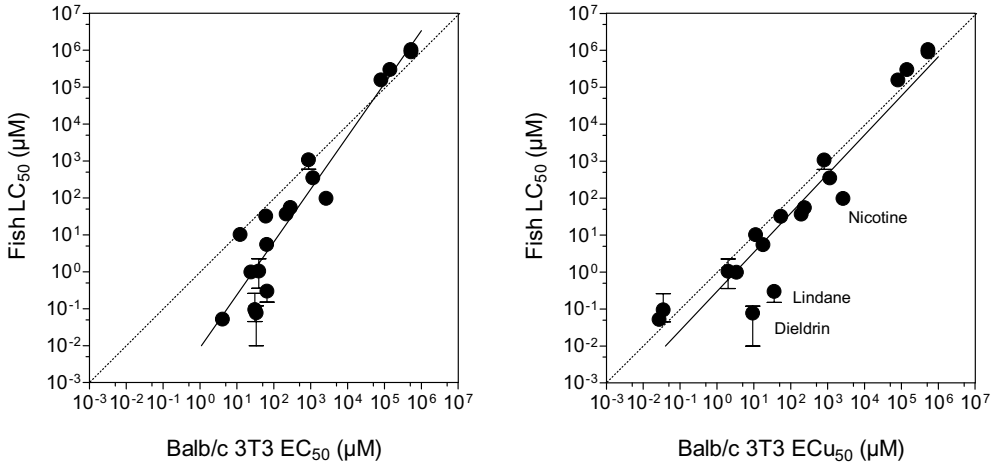


Fig 4: Comparison of the aquatic acute toxic concentrations of chemicals to fathead minnow (*Pimephales promelas*) with their nominal cytotoxic (left) and free cytotoxic (right) concentrations measured *in vitro* using the mouse Balb/c 3T3 cell proliferation inhibition assay. Data are from Gülden and Seibert (2005). The dotted lines indicate 1:1 correlation. The solid lines represent the linear regression lines:
left: $\log LC_{50} = -2.06 + 1.43 \times \log EC_{50}$, $r^2 = 0.911$
right: $\log LC_{50} = -0.52 + 1.06 \times \log ECu_{50}$, $r^2 = 0.901$

(5) Otherwise, the specific binding to albumin (B) has to be measured and the free effective concentration is given by:

a) If $K_{ow} < 1/V'_L$

$$ECu_{50} = EC_{50} - B \bullet P \quad (5)$$

b) If $K_{ow} \geq 1/V'_L$

$$ECu_{50} = \frac{EC_{50} - B \bullet P}{1 + K_{ow} \bullet V'_L} \quad (2)$$

Application example 2: Prediction of human serum concentrations which are equivalent to cytotoxic concentrations *in vitro*

The albumin and lipid contents of animal and human serum are considerably higher than those of the extracellular media in *in vitro* tests systems. Consequently, the availability of lipophilic and albumin bound chemicals in serum must be lower than *in vitro*. To attain the same free concentration of a chemical, in such cases, higher nominal concentrations have to be achieved. On the basis of the equilibrium distribution model described above, we have developed the following algorithm to extrapolate the serum concentrations (EC_{serum}) of chemicals which are equivalent to the *in vitro* effective concentrations (Gülden and Seibert, 2003):

$$EC_{serum} = (EC_{50} - B \bullet P_{invitro}) \frac{1 + K_{ow} \bullet V'_{L,\,serum}}{1 + K_{ow} \bullet V'_{L,\,invitro}} + B \bullet P_{serum} \quad (6)$$

where $V'_{L,invitro}$ and $V'_{L,serum}$ are the relative lipid volumes, and $P_{invitro}$ and P_{serum} the albumin concentrations in the medium of the *in vitro* system and in serum, respectively.

This extrapolation model was applied to cytotoxic concentrations determined *in vitro* with the Balb/c 3T3 proliferation inhibition test in order to estimate equivalent human serum concentrations. In total 33 chemicals covering a wide range of cytotoxic potency and lipophilicity were investigated. Octanol/water partition coefficients were available for all of these chemicals. The specific binding to albumin was determined by EC_{50}-measurements in the presence of different albumin concentrations (Gülden et al., 2002, 2003).

The concentration of albumin in human serum (~ 600 µM) is about 33-times higher than in the medium of the Balb/c 3T3 test system used, and the relative volume of lipids in serum, with a mean of $6x10^{-3}$ l/l, is about 60-times higher.

The application of the extrapolation model revealed that for 13 of the 33 compounds the equivalent serum concentrations were equal to the nominal EC_{50}-values obtained *in vitro* (fig. 5).

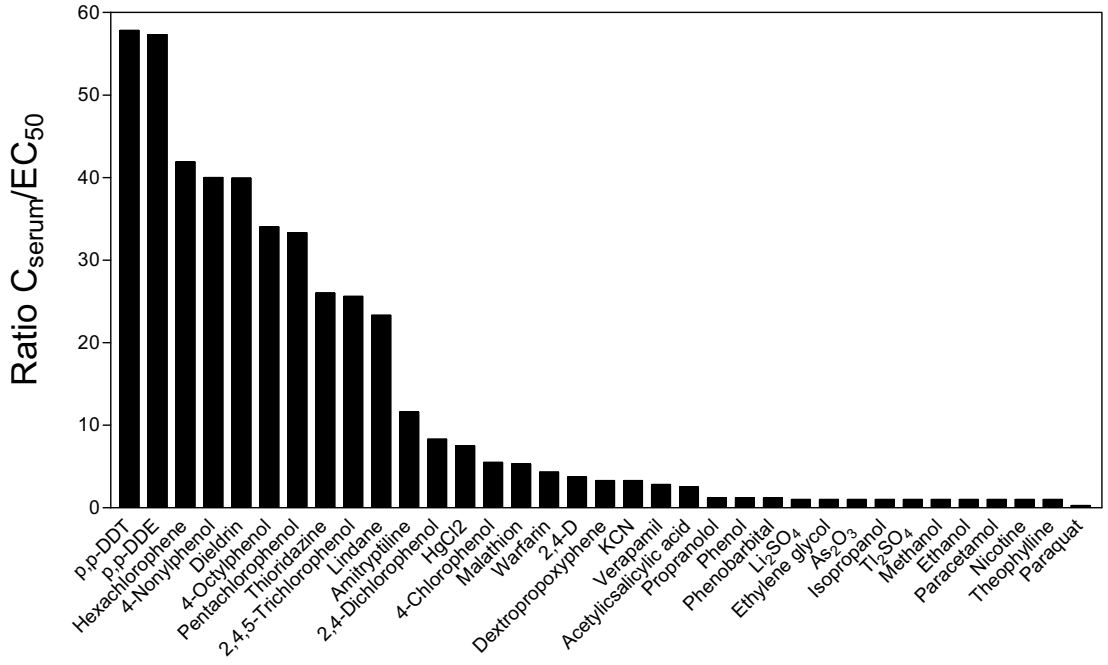


Fig. 5: Ratios of the equivalent human serum concentrations and the nominal cytotoxic concentrations *in vitro* determined with the mouse Balb/c 3T3 cell proliferation inhibition assay. The equivalent serum concentrations (C_{serum}) were calculated from the nominal cytotoxic concentrations *in vitro* (EC_{50}) using the extrapolation model Eq. (6), albumin concentrations *in vitro* and in serum of 18 µM and 600 µM, and lipid volume fractions *in vitro* and in serum of 10^{-4} and $6x10^{-3}$ l/l, respectively.

Neither lipophilicity nor albumin binding capability of these chemicals were high enough to affect their availability in the *in vitro* system or in serum. The equivalent human serum concentrations of the other chemicals were substantially higher, by factors of 2.5 to 58, compared to the nominal EC_{50}-values. The lipophilicity and/or protein binding capacity of these chemicals were at least high enough to affect their bioavailability in serum.

A further analysis of the results indicated that differences in the availability of chemicals become more frequent with increasing potency. In case of chemicals with low toxic potency neither lipophilicity nor albumin binding is high enough to have relevant influence on their availability. Neglecting differences in the availability of chemicals *in vitro* and *in vivo*, thus, will systematically bias the correlation between *in vitro* and *in vivo* toxic potencies. First, the slopes of the regression lines describing the correlations between the log transformed potencies must clearly deviate from the 1:1 correlation and be lower than 1. Second, the scatter of data points around the regression line must increase with increasing *in vitro* potencies. Both these phenomena are seen when acute toxic potencies (e.g. LD_{50}) are correlated with cytotoxic potencies determined *in vitro* (Garle et al., 1994; Halle, 2003).

Recently, the validity of this extrapolation model was evaluated. The results indicate that the extrapolation model is well suited to predict equivalent concentrations in serum from *in vitro* effective concentrations (Gülden et al., 2006).

Practical strategy for extrapolating effective serum concentrations from *in vitro* test results

From these results and theoretical considerations on the impact of protein binding and partitioning into lipids on the bioavailability of chemicals in serum and in *in vitro* test systems, some pragmatic rules for the extrapolation of equivalent chemical concentrations in human serum can be deduced (see Gülden and Seibert, 2003). These rules are not restricted to cytotoxic potencies but can be applied to any biological activity of chemicals determined *in vitro*. The premise is that the distribution relevant composition (V'_L = lipid volume fraction, P = albumin concentration) of both the *in vitro* system and the serum concerned and the octanol/water partition coefficients of the chemicals are known.

1) If the octanol/water partition coefficient of the chemical is lower than the reciprocal of the relative lipid volume in serum ($Kow < 1/V'_{L,serum}$), partitioning into lipids is not likely to affect the availability of the chemical in serum or *in vitro*. In case of human serum, with a relative lipid volume of $V'_{L,serum} \leq 10^{-2}$ l/l, this is the case if $K_{ow} \leq 10^2$.

2) If the EC_{50}-value of a chemical measured *in vitro*, either in the absence or the presence of comparatively low serum protein concentrations (i.e. up to 20% serum in the medium), is higher than the albumin concentration in the serum concerned for extrapolation ($EC_{50} > P$), a quantitative impact of protein binding on the availability of the chemical is not to be expected. For human serum this is the case if the EC_{50} is higher than 600 µM. Then, the equivalent serum concentration (EC_{serum}) is given by the following equations:

a) If $K_{ow} < 1/V'_{L,serum}$

$$EC_{serum} = EC_{50} \quad (7)$$

b) If $K_{ow} \geq 1/V'_{L,serum}$

$$EC_{serum} = EC_{50} \frac{1 + K_{ow} \bullet V'_{L,serum}}{1 + K_{ow} \bullet V'_{L,invitro}} \quad (8)$$

3) Otherwise, the specific binding (B) to albumin has to be measured, and the equivalent serum concentration is given by

a) If $K_{ow} < 1/V'_{L,serum}$

$$EC_{serum} = EC_{50} + B(P_{serum} - P_{invitro}) \quad (9)$$

b) If $K_{ow} \geq 1/V'_{L,serum}$

$$EC_{serum} = (EC_{50} - B \bullet P_{invitro}) \frac{1 + K_{ow} \bullet V'_{L,serum}}{1 + K_{ow} \bullet V'_{L,invitro}} + B \bullet P_{serum} \quad (6)$$

Conclusion

In conclusion, taking into consideration the elementary differences in the bioavailability of chemicals between *in vitro* and *in vivo* systems improves the usage of *in vitro* generated toxicity data in the process of toxicological hazard and risk assessment. The concept of equivalent exposure for *in vitro-in vivo* extrapolation of toxic potencies takes such differences in the availability into account. Applying this concept to predict acute fish toxicity and toxic concentrations in human serum from *in vitro* cytotoxic concentrations demonstrates the practical value of this approach.

References

Bohets H. H., Nouwen E. J., De Broe M. E. and Dierickx P. J. (1994). Effects of foetal calf serum on cell viability, cytotoxicity and detoxification in the two kidney derived cell lines LLC-PK1 and MDCK. *Toxicol. in Vitro 8*, 559-561.

Castano, A., Bols, N., Braunbeck, T. et al. (2003). The use of fish cells in ecotoxicology. Report and recommendations of the ECVAM workshop 47. *ATLA 31*, 317-351.

Castano, A. and Gomez-Lechon, M. (2005). Comparison of basal cytotoxicity data between mammalian and fish cell lines: A literature survey. *Toxicol. in Vitro 19*, 695-705.

Finlay, G. J. and Baguley, B. C. (2000). Effects of protein binding on the in vitro activity of antitumour acridine derivatives and related anticancer drugs. *Cancer Chemother. Pharmacol. 45*, 417-422.

Garle, M. J., Fentem, J. H. and Fry, J. R. (1994). In vitro cytotoxicity tests for the prediction of acute toxicity in vivo.

Toxicol. in Vitro 8, 1303-1312
Gülden, M. and Seibert, H. (1997). Influence of protein binding and lipophilicity on the distribution of compounds in in vitro systems. *Toxicol. in vitro 11*, 479-483.
Gülden, M., Mörchel, S. and Seibert, H. (2001). Factors influencing nominal effective concentrations of chemical compounds in vitro: cell concentration. *Toxicol. in Vitro 15*, 233-243.
Gülden, M., Mörchel, S., Tahan, S. and Seibert, H. (2002). Impact of protein binding on the availability and cytotoxic potency of organochlorine pesticides and chlorophenols in vitro. *Toxicology 175*, 201-213.
Gülden, M., Mörchel, S. and Seibert, H. (2003). Serum albumin binding at cytotoxic concentrations of chemicals as determined with a cell proliferation assay. *Toxicol. Lett. 173 (3)*, 159-168.
Gülden, M. and Seibert, H. (2003). In vitro-in vivo extrapolation: estimation of human serum concentrations of chemicals equivalent to cytotoxic concentrations in vitro. *Toxicology 189*, 211-222.
Gülden, M., Mörchel, S. and Seibert, H. (2005). Comparison of mammalian and fish cell line cytotoxicity: impact of endpoint and exposure duration. *Aquat. Toxicol. 71*, 229-236.
Gülden, M. and Seibert, H. (2005). Impact of bioavailability on the correlation between in vitro cytotoxic and in vivo acute fish toxic concentrations of chemicals. *Aquat. Toxicol. 72*, 327-337.
Gülden, M., Dierickx, P. and Seibert, H. (2006). Validation of a prediction model for estimating serum concentrations of chemicals which are equivalent to toxic concentrations in vitro. *Toxicol. in Vitro,* in press.
Halle, W. (2002). The registry of cytotoxicty: Toxicity testing in cell cultures to predict acute toxicity (LD_{50}) and to reduce testing in animals. *ATLA 31*, 89-198.
Heringa, M. B., Schreurs, R .H. M. M., Busser, F. et al. (2004). Toward more useful in vitro toxicity data with measured free concentrations. *Environ. Sci. & Technol. 38*, 6263-6270.
Hestermann E. V., Stegeman J. J. and Hahn M. E. (2000). Serum alters the uptake and relative potencies of halogenated aromatic hydrogencarbons in cell culture bioassays. *Toxicol. Sci. 53*, 316-325.
Segner, H. (1998). *Toxicology 189*, 211-222.
Segner, H. (2004). Cytotoxicity assays with fish cells as an alternative to the acute lethality test with fish. *ATLA 32*, 375-382.
Segner, H. and Lenz, D. (1993). Cytotoxicity assays with the rainbow trout R1 cell line. *Toxicol. in Vitro 7*, 537-540.
Seibert, H., Gülden, M., and Voss, J. U. (1994) An in vitro toxicity testing strategy for the classification and labelling of chemicals according to their potential acute lethal potency. *Toxicology in Vitro 8*, 847-850.
Seibert, H., Balls, M., Fentem, J. H. et al. (1996). Acute toxicity testing in vitro and the classification and labelling of chemicals. The report and recommendations of ECVAM workshop 16. *ATLA 24*, 499-510.
Seibert, H., Mörchel, S. and Gülden, M. (2002). Factors influencing nominal effective concentrations of chemical compounds in vitro: serum albumin concentration. *Toxicol. in Vitro 16*, 289-297.
Vaes, W. H. J., Urrestarazu, Ramos E., Hamwijk C. et al. (1997). Solid phase microextraction as a tool to determine membrane/water partition coefficients and bioavailable concentrations in in vitro systems. *Chem. Res. Toxicol. 10*, 1067-1072.
Worth, A. P. and Balls, M. (eds.) (2002). Alternative (Non-animal) Methods for Chemicals Testing: Current Status and Future Prospects. A Report prepared by ECVAM and the ECVAM Working Group on Chemicals. *ATLA 30, Supplement 1*.

Correspondence to
Dr. Michael Gülden
Institute of Toxicology and Pharmacology for Natural Scientists
University Medical School Schleswig-Holstein
Campus Kiel
Brunswiker Str. 10
24105 Kiel
Germany
e-mail: guelden@toxi.uni-kiel.de

Session 5.03
Progress and needs for developing and validating alternatives for dermal sensitisation testing

Dendritic Cell Research Projects

Pierre Aeby[2], David Basketter[1] (Chairman), Walter Diembeck[3], Dietmar Eschrich[4], Frank Gerberick[5], Hiroshi Itagaki[6], Ian Kimber[7], Béatrice Le Varlet[8], Irene Manou[9], Marc Paye[10], Françoise Rousset[11], Joanna Rowland[12] and Hitoshi Sakaguchi[13]

[1]Unilever Safety and Environmental Assurance Centre, Sharnbrook, UK; [2]Cosmital (Wella), TOX, Marly, Switzerland; [3]Beiersdorf, Hamburg, Germany; [4]Phenion/Henkel, Frankfurt, Germany; [5]The Procter & Gamble Company, Cincinnati, OH, USA; [6]Shiseido Research Centre, Yokohama-Shi, Japan; [7]Syngenta Central Toxicology Laboratory, Macclesfield, UK; [8]LVMH-Recherche, Saint Jean de Braye, France; [9]COLIPA, Brussels, Belgium; [10]Colgate, R&D, Milmort, Belgium; [11]L'Oreal, Advanced Research, Aulnay sous Bois, France; [12]GlaxoSmithKline, Consumer Healthcare, R&D, Weybridge, UK; [13]Kao Corporation, Tochigi, Japan

Summary

The Colipa Skin Tolerance Task Force (STTF) is playing an active role in promoting refined methods for in vivo sensitisation tests. Dendritic cells (DCs), whose central role during the induction phase of skin sensitisation is well documented, were perceived as promising in vitro test systems. DC research projects supported by STTF evaluated the relevance of IL-1β mRNA expression or changes in gene expression using gene microarrays for predicting potential sensitisers. Myeloid cell lines are being evaluated as surrogate DCs. Intracellular signal transduction pathways are being explored in order to understand the underlying biological mechanisms.

Keywords: dendritic cells, contact sensitisation, hapten, in vitro model

Introduction

The sensitising potential of chemicals is usually identified on the basis of animal studies. In addition to increasing public concern regarding the use of animal testing for the toxicological evaluation of new chemicals, the adoption of the 7th Amendment to the Cosmetics Directive in Europe will result in animal testing of ingredients used in cosmetics being subjected to severe restrictions by 2009 which will be tightened by 2013. The development of *in vitro* models for predicting the sensitising potential of new chemicals is therefore an obvious need. *In vitro* sensitisation tests will have to reflect the complex interactions of a chemical with the different compartments of the immune system. COLIPA STTF is therefore supporting a range of research projects reflecting our current understanding of the molecular and cellular events occurring during the sensitisation process from aspects of chemistry, peptide binding and skin metabolism to intracellular signalling pathways induced by allergens (see fig. 1).

The Colipa dendritic cell research projects

Dendritic cells (DCs), whose central role during the induction phase of skin sensitisation is well documented (Enk et al., 1993; Lepoittevin et al., 1998; Aiba et al., 1997; Smith and Hotchkiss, 2001), were rapidly perceived as promising *in vitro* test systems. The publication of a paper describing specific *in vitro* upregulation of IL-1β mRNA in skin DCs exposed to sensitisers (Reutter and Jaeger, 1997) convinced STTF to initiate a research project aimed at evaluating the relevance of IL-1β mRNA expression in

cultures of human DCs for predicting potential sensitisers. In 1998, the STTF decided to support a research project to evaluate the relevance of IL-1β mRNA expression by cultured dendritic cells as an endpoint for predicting potential skin sensitisers. Dendritic cells were cultured from peripheral blood drawn from human volunteers. Although donor specific variation in the phenotype of DCs was observed, all blood-derived DCs constitutively expressed mRNA for IL-1β, IL-6 and IL-18. A significant increase (2-3 fold) in IL-1β mRNA expression was observed with DCs derived from 4 out of 9 donors after *in vitro* exposure to the known contact sensitiser 2,4-dinitrofluorobenzene (DNFB). The variation in DNFB-induced up-regulation of IL-1β expression appeared to be donor dependent. In no instance did treatment of DCs with the non-sensitising skin irritant sodium lauryl sulphate (SLS) increase IL-1β mRNA expression. Exposure to DNFB was largely without effect on constitutive IL-6 or IL-18 mRNA expression. However, although apparently selective for skin sensitising chemicals, the induced changes in IL-1β expression were modest and it is not known whether less potent sensitising chemicals would have the same ability to influence IL-1β mRNA levels (Pichowski et al., 2000).

This work demonstrated that there is some basis for evaluating IL-1β, and certain other genes, based upon what was understood of their mechanistic relevance in the development of skin sensitisation. However, they may not necessarily represent the most appropriate markers for the identification of skin sensitisers and skin irritants.

The STTF thus decided to support a further project proposal making use of the recent availability of DNA arrays, which permit characterisation of induced changes in gene expression in a more holistic fashion. Using this technology, gene expression changes in peripheral blood-derived dendritic cells were analysed following exposure to a contact allergen. Briefly, peripheral blood-derived DCs were exposed for 24 h to either 1 mM or 5 mM dinitrobenzenesulphonic acid (DNBS). Changes in gene expression were analysed using Affymetrix© gene chips. 1 mM and 5 mM DNBS induced 173, respectively 1249 significant gene changes. Many of the gene expression changes were found

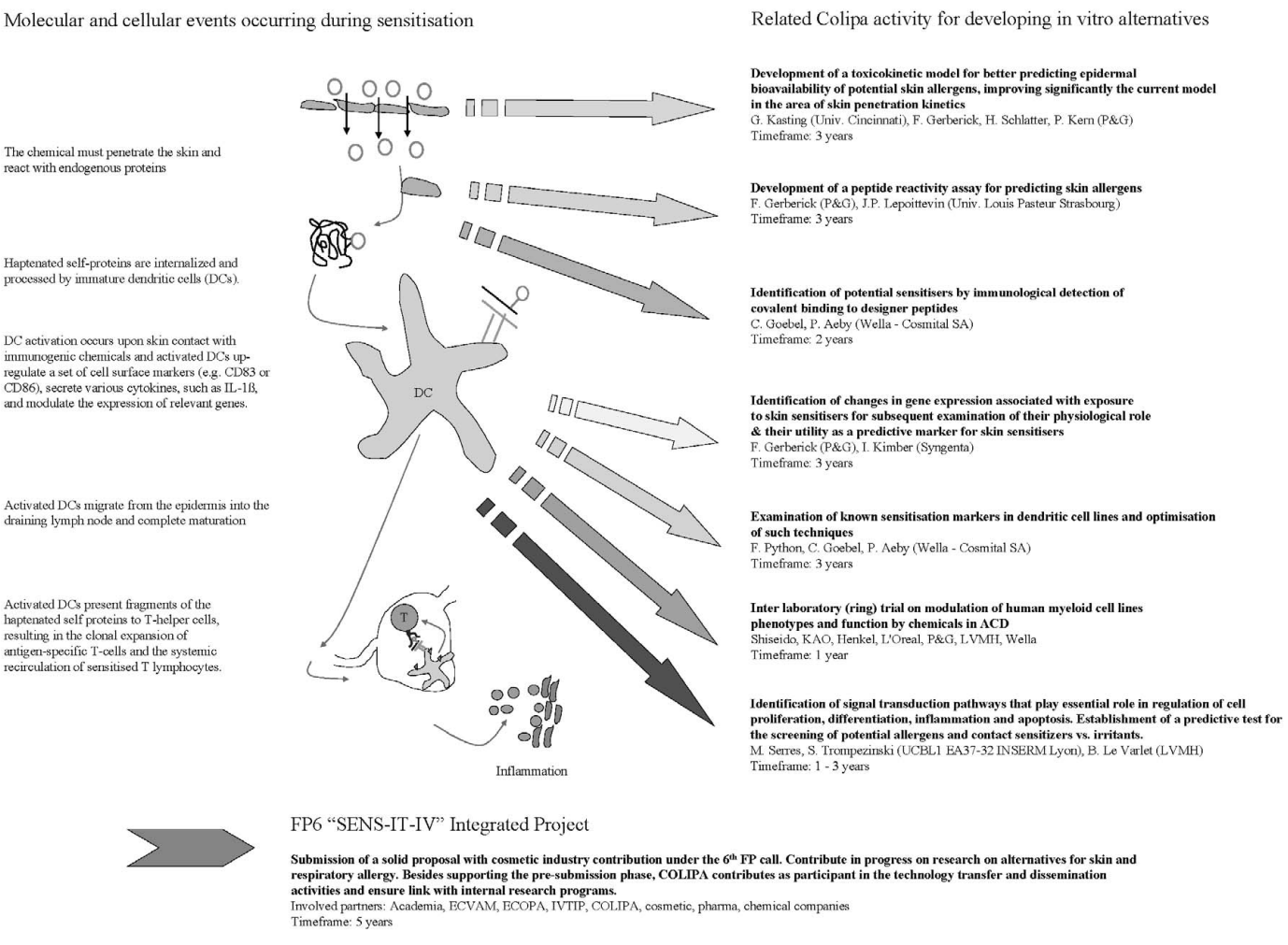


Fig. 1: Molecular and cellular events occurring during sensitisation and related Colipa activity for developing *in vitro* alternatives.

to be induced only by exposure to DNBS and not by exposure to the non-allergen benzenesulfonic acid (Ryan et al., 2004). These data, in part, provided support of the theory that exposure to contact allergens can induce the maturation of DCs. Additionally, some of the transcript changes identified through this approach may be shown to be sufficiently robust, sensitive and selective, such that they are suitable for use in the development of an *in vitro* predictive assay (Ryan et al., 2004).

There was also a need for a source of homogeneous and reproducible DCs. STTF thus initiated a research project concerning the analysis of the modulation of human myeloid cell lines' phenotypes and functions by chemicals in 2002.

Three human myeloid cell lines, U-937, THP-1 and KG-1, were evaluated as a replacement of human blood derived DCs. The cytokine environment (granulocyte-macrophage colony-stimulating factor, interleukin-4 (IL-4), stem cell factor, transforming growth factor-β1 and tumour necrosis factor-α (TNF-α)) was first optimised for the induction of differentiation into DC-like cells. TNF-α was used to determine the maximum extent and the kinetics of CD86 modulation. The results indicated that among the three tested lines, IL-4 pre-treated U-937 cells provided the optimal CD86 expression range. In parallel, IL-1β and IL-8 mRNA expression was measured by real time RT-PCR. The modulation of the three markers CD86, IL-1β and IL-8 allowed discrimination of sensitisers from non-sensitisers. Results obtained for sensitisers indicate that a combination of at least two of these markers at more than one time point is needed to establish a reliable classification system (Python et al., 2004).

DCs or DC cell lines seem to be specifically activated by sensitising chemicals. In order to better understand the underlying molecular mechanisms, the STTF decided in 2004 to support a project proposal concerning the exploration of intracellular signal transduction pathways in activated dendritic cells.

Experiments were performed on immature human monocyte-derived dendritic cells, induced to mature by LPS, TNF-α or nickel treatment. The pathways of the main kinases known from literature to be involved in intracellular signalling were investigated by use of specific inhibitors at each step of the phosphorylation cascade. Preliminary results showed that CD86 expression induced by DC maturation is influenced by some of these inhibitors, suggesting an essential role of the respective kinases (Trompezinski et al., 2005).

The Colipa Skin Tolerance Task Force (STTF) is playing a very active role in promoting research on *in vitro* alternatives to *in vivo* sensitisation tests. Its DC research programme is a key part of an integrated effort to develop alternatives to animal tests (see fig. 1).

References

Aiba, S., Terunuma, A., Manome, H., and Tagami, H. (1997). Dendritic cells differently respond to haptens and irritants by their production of cytokines and expression of co-stimulatory molecules. *Eur. J. Immunol. 27*, 3031-3038.

Enk, A. H., Angeloni, V. L, Udey, M. C. and Katz, S. I. (1993). An essential role for Langerhans cell-derived IL-1 beta in the initiation of primary immune responses in skin. *J. Immunol. 150*, 3698.

Lepoittevin, J. P., Basketter, D. A., Goossens, A. and Karlberg, A-T. (ed.) (1998). *Allergic Contact Dermatitis – The Molecular Basis*. Berlin: Springer-Verlag.

Pichowski, J. S., Cumberbatch, M., Dearman, R. J. et al. (2000). Investigation of induced changes in interleukin 1ß mRNA expression by cultured human dendritic cells as an in vitro approach to skin sensitization testing. *Toxicol. In vitro 14,* 351-360.

Python, F., Goebel, C. and Aeby, P. (2004). Evaluation and optimization of an in vitro culture system for the identification of contact allergenes, Abstract from the ICTX 2004. *Toxicol. Appl. Pharmacol. 197(3)*, 182.

Reutter, K. and Jaeger, D. (1997). In vitro model for contact sensitization: II. Induction of IL-1ß mRNA in human blood-derived dendritic cells by contact sensitizers. *Toxicol. In vitro 11*, 619-626.

Ryan, C. A., Gildea, L. A., Hulette, B. C. et al. (2004). Gene expression changes in peripheral blood-derived dendritic cells following exposure to a contact allergen. *Toxicol. Lett. 150(3)*, 301-316.

Smith, C. K. and Hotchkiss, S. A. M. (eds.) (2001). *Allergic Contact Dermatitis – Chemical and Metabolic Mechanisms. Chapter 1: Allergic Contact Dermatitis to Small Molecule Xenobiotics.* London and New York: Taylor & Francis.

Trompezinski, S., Denoeud, J., Zabawski B. et al. (2005). NF-κB, PI3K and JNK are essential signaling pathways implicated in human dendritic cells maturation after contact with allergens. Abstract from the 5th World Congress on Alternatives & Animal Use in the Life Sciences, Berlin 2005.

Correspondence to

Dr. Pierre Aeby
Cosmital SA
Route de Chesalles 21
1723 Marly, Switzerland
e-mail: aeby.p@pg.com

Hapten-Protein Binding: What Do We Know?

Maja Divkovic
Safety & Environmental Assurance Centre, Unilever Colworth, Sharnbrook, Bedfordshire, UK

Summary
The chemical modification of self skin proteins (protein haptenation) and formation of macromolecular immunogens is one of the key molecular events in skin sensitisation. This widely accepted concept can be used to explain the sensitising capacity of many known skin sensitisers and could be used in in vitro assays to predict the sensitisation potential of a chemical. Here we review our current knowledge, highlight the gaps in our understanding of protein haptenation and discuss how we can use such knowledge in the development of novel alternative in vitro approaches for predicting skin sensitisation potential in the future.

Keywords: skin sensitisation, hapten-protein binding, peptide binding, in vitro assay

Introduction

Allergic contact dermatitis (ACD) is a delayed (type IV) hypersensitivity reaction to an exogenous chemical mediated by T-cell related processes (Girolomoni et al., 2004; Rustemeyer et al., 2004; Coombs and Gell, 1975). However, validated *in vitro* alternatives for assessment of skin sensitisation potential and estimation of relative potency are not available to date, so these assessments are based on *in vivo* methods, such as the murine local lymph node assay (LLNA) (Basketter et al., 2001). Improving our understanding of the cellular and molecular mechanisms of the sensitisation process could result in novel opportunities for development of alternative *in vitro* methods for assessing skin sensitisation hazard and relative potency of chemicals.

Several approaches are being investigated to expand our understanding of the sensitisation process (reviewed in Casati et al., 2005). Additionally, there are a number of other published cell culture studies where responses to treatment with sensitisers have been investigated, including keratinocyte cultures and co-cultures of dendritic cells with T cells (Krasteva et al., 1996; Rougier et al., 2000; Rougier et al., 1998; Guironnet et al., 2000). Signal transduction pathways in dendritic cells, such as mitogen-activated protein kinases (MAPKs) and NFκB pathways have also been a subject of several investigations in search of reliable markers (Aiba et al., 2003; Boisleve et al., 2005).

Exploring the mechanisms of hapten-protein binding in the early stages of the skin sensitisation process is based on the hypothesis that upon skin absorption, only protein-reactive chemicals (or those that can be metabolically or chemically converted to protein-reactive species) are able to act as skin sensitisers, and that they do this through a process of protein haptenation (Dupuis and Benezra, 1982; Landsteiner and Jacobs, 1935).

The "covalent binding" concept

The landmark paper by Landsteiner and Jacobs proposed that the sensitising chemicals are too small to be recognised by the classical immunological mechanisms. Therefore they need to be bound to protein in order to elicit an immune response. Unless already a protein-reactive molecule, a sensitiser may be chemically or metabolically activated prior to or upon cutaneous absorption.

Generally, reactions of chemicals with proteins lead to the formation of bonds of different strengths. Weak interactions, such as hydrophobic, dipolar (including hydrogen bonds) and ionic bonds involve energies of up to 50 kJ/mol, whereas strong interactions, such as covalent and co-ordination bonds, involve energies ranging from 200 to 420 kJ/mol (Lepoittevin, 2001). Despite the suggestions that non-covalent modes of hapten-protein association or modification of the normal self protein processing pathways and consequent creation of "cryptic" epitopes could lead to an immune response (Pichler, 2001), evidence to date indicates that the interaction between the hapten and the protein must result in the formation of a strong bond.

Over the last years, the methodology (immunochemistry, NMR and mass spectrometry) applied to characterise protein-hapten binding has improved substantially, enabling us to investigate the specificity as well as the extent of protein binding using a variety of skin sensitisers. Covalent binding is now widely accepted as a result of investigations of numerous examples of skin sensitisers and their ability to covalently modify proteins. Hapten-protein binding studies have achieved important milestones in our understanding of protein haptenation mechanisms, despite not always being related to skin sensitisation (Ahmed et al., 2005; Baker et al., 1998; Conduah Birt et al., 1998; Nerland et al., 2003; Person et al., 2003; Tracey and Shuker, 1997; Walker, 1976; Fabrizi et al., 2003).

Considering the chemicals

Chemicals may react with many different skin proteins at different amino acid sites, but in general protein molecules are rich in nucleophiles and the sensitising chemicals are reactive electrophiles or can be metabolically converted to such. The potential reaction mechanisms that lead to skin sensitisation have been discussed by many authors (Basketter et al., 1995; Smith and Hotchkiss, 2001) but very few of them have been proven experimentally.

Considering the target

The main targets of small molecule electrophiles are amino acid side chains with nucleophilic properties. Nucleophiles either have atoms containing one or more unshared pairs of electrons or are negatively charged ions. The strongest potential nucleophiles in proteins, apart from the N-terminal amino group, are the lysine ε-amino group, the cysteine sulphhydryl group and the histidine imidazole group. The ability of amino acid side chains to react with electrophilic chemicals is largely dependent on the degree of ionisation, bearing in mind that the skin pH ranges from 5.5 on the *stratum corneum* surface to physiological 7.4 in the epidermis and dermis (Smith and Hotchkiss, 2001). The ability of a chemical to react with a nucleophile may be hindered or enhanced by the nucleophile's position in a 3D protein environment.

Skin as a complex heterogenous tissue expresses a large number of proteins. Over 2000 proteins have been separated in cultured human keratinocytes, but approximately only a third of those could be identified using proteomics techniques (Celis et al., 1998). It is logical to assume that a protein reactive chemical will modify any available nucleophile to some extent given suitable conditions. A recent study investigating the binding of cinnamaldehyde to human skin homogenates showed that this moderate sensitiser was bound to a broad range of proteins in the sample and there did not appear to be any specific targeting to any particular proteins (Elahi et al., 2004). In the absence of a target skin protein per se, *in vitro* investigations have been limited to the use of model proteins or peptides to explore general chemistries.

Protein/peptide models for studying peptide haptenation

A number of different models have been used to study protein haptenation. Human serum albumin (HSA) is often the model of choice for protein binding assays. This is a well characterised protein and around 40% of extravascular HSA is located in the skin (Peters, 1996). Several key HSA residues have been shown to be selectively and covalently modified by chemicals, such as Cys34 (Yasuzawa and Tomer, 1997), Lys 190 (Bohney et al., 1992), Lys 199 (Walker, 1976; Bertucci et al., 1995; Meschkat et al., 2001), His 9, His 146, His 338 (Alvarez-Sanchez et al., 2004) and Arg 410 (Ahmed et al., 2005). HSA has 17 pairs of cysteines involved in disulphide linkages and only one free Cys residue (Cys34). Similarly, there is only one Trp residue on HSA. Other macromolecules with different properties and amino acid composition may reveal different mechanisms of protein haptenation for the same chemical. The question about the immunological relevance of specific residue modifications in macromolecules remains. For example, it has been suggested that a tolerising effect may be due to exclusive modification of sulphydryl groups (Parker et al., 1983).

Peptides with sequences analogous to a part of human proteins are also used to assess chemical reactivity. DS3 peptide, which has a sequence analogous to the N-terminal part of the human globin (sequence VLSPADKTNWGHEYRMFCQIG), was used to investigate binding of 4-chlorobenzenediazonium hexafluorophosphate (Tracey and Shuker, 1997), acetaldehyde (Conduah Birt et al., 1998) and 5-chloro-2-methylisothiazol-3-one (Alvarez-Sanchez et al., 2004). Glycine apart, this peptide contains one residue of each of the commonly occurring amino acids, with Cys residue of the peptide often carboxymethylated or simply omitted from synthesis to avoid peptide dimerisation in reactivity studies. Similarly, the synthetic peptide PEPAK-SAPAPKKGSKKAVTKAQK was used to investigate reactions of phosgene (Fabrizi et al., 2003).

Synthetic peptides unrelated to any proteins have also been used as models such as N-acetylglycyllysine O-methyl ester (AcGKOMe), which was used to investigate the complex reactions of 2-alkenals (Baker et al., 1998). Similarly, a synthetic peptide PHCKRM was used to investigate the binding of 1,4-benzoquinone (Ahlfors et al., 2003), 4-*t*-butyl-1,2-benzoquinone (Hanson et al., 1999) and two metabolites of a prohapten (5R)-5-isopropenyl-2-methyl-1-methylene-2-cyclohexene (Nilsson et al., 2005; Ahlfors et al., 2003).

The major problem with using small peptides is the further step away from the biological conditions in the skin, as the potentially crucial influence of the 3-dimensional protein environment is not represented. However, model peptides have their place in such investigations, proving explanations for chemistries that may currently only be theoretical for many sensitisers and revealing new potential chemistries (Dupuis and Benezra, 1982; Smith and Hotchkiss, 2001). A further advantage is the simpler analytical process for short peptides and thus the potential for a medium or high throughput assay to be developed as well as the ability to use quantification.

Some investigators have opted to use model nucleophiles such as butylamine, imidazole and propanethiol representing lysine, histidine and cysteine (respectively) to investigate reactivity of chemicals towards proteins (Franot et al., 1994; Meschkat et al., 2001; Alvarez-Sanchez et al., 2003).

Regardless of the protein/peptide models used, these investigations provide insights into often complicated chemistries involved in modification of proteins/peptides. Particularly useful are comparative studies of chemically related sensitisers with differing potencies as illustrated in experiments with hex-1-ene and hexane-1,3-sultones, a strong and moderate sensitiser, respectively (Meschkat et al., 2001) and butyrolactones (Franot et al., 1994).

However, depending on the models used, different conclusions can be made about the chemical reactivity of certain chemical entities. This is illustrated in the series of papers about reactivity of 5-chloro-2-methyl-4-isothiazolin-3-one (MCI) with

different models. MCI is an extreme sensitiser (Basketter et al., 1999; Gerberick et al., 2004) and a constituent of Kathon CG (Collier et al., 1990a; Collier et al., 1990b), a microbiocide used as a preservative in skin care products. The reactions of MCI have been investigated with model nucleophiles (Alvarez-Sanchez et al., 2003), model peptide, GSH (Alvarez-Sanchez et al., 2004) and HSA (Alvarez-Sanchez et al., 2004). The proposed substitution reaction was confirmed by ^{13}C NMR and shown to take place with imidazole (His in model peptide and HSA studies), whereas reaction with butylamine (Lys in model peptide and HSA studies) required a prior activation with thiol, which gave adducts of amide and thioamide type.

Similarly, the chemistry of benzoquinone also depends on the model used. The study which utilised model peptide PHCKRM showed benzoquinone reacted with Cys exclusively (Hanson et al., 1999; Ahlfors et al., 2003), whereas the study which utilised cytochrome *c* as a model protein, showed benzoquinone reacted with two adjacent Lys residues resulting in cyclised diquinone product (Person et al., 2003). Which of the two modifications represents the immunogenic entity remains to be shown.

In a study unrelated to skin sensitisation, another model peptide (N-acetylglycyllysine *O*-methyl ester, (AcGKOMe)) was used in a study to examine mechanistic aspects of modification of the ε-amino group of lysine by 2-alkenals (Baker et al., 1998). The reaction products were characterised using a combination of ^{1}H and ^{13}C NMR, HPLC separation, FAB MS, LC-MS and MALDI-MS. The results indicate that one ε-amino group of lysine can react with two alkenals to form a dihyropyridine ring, and this can further react to form pyridinium moieties or stable pyridinium cross-links. Such complex adducts to proteins of more than one molecule of the reacting chemical are not uncommon. Creation of such elaborate adducts could result in a more vigorous immune response.

Conclusions and future directions

The ultimate goals within the skin sensitisation field are to develop better *in vitro/in silico* tools for screening of chemicals for skin sensitisation hazard in early product development and to devise alternative approaches to replace *in vivo* assays for predicting sensitisation hazard and potency in man (Divkovic et al., 2003). Recent advances in understanding the molecular basis of skin sensitisation and the associated chemistries are providing more evidence for the covalent protein haptenation theory but are also revealing additional mechanistic complexity. Further insights have become available in terms of the cellular localisation of hapten-protein binding and further studies in this area would be informative. The selectivity of haptenation for cellular and extracellular proteins with different types of chemicals (contact and respiratory allergens) was recently investigated (Hopkins et al., 2005). When incubated with cells and serum together, contact sensitisers were found to selectively bind to cellular proteins, as opposed to respiratory sensitisers which selectively bound to serum proteins. It would be interesting to explore the cellular locality of protein haptenation further (Smith and Hotchkiss, 2001). Additionally, advances in cutaneous functional proteomics are providing critical data in our understanding of skin sensitisation (Huang et al., 2005). This may help determine the immunologically relevant proteins which are targets for sensitising molecules. Better predictive assays could be designed if particular types of modifications were found more immunogenic than others. Equally, better model proteins/peptides could be chosen if we found out what exactly is presented to the naïve T cell associated with MHC II. It is at least theoretically possible to isolate the MHC II molecules and elute the peptides attached via hydrogen bonds using acid elution (Hayden et al., 1996; Lippolis et al., 2002; Sharif et al., 2003; Suri et al., 2003; Park et al., 2003). Sophisticated ultra sensitive mass spectrometry techniques would allow unambiguous identification of the peptide sequence and any modifications present. If covalently modified peptides are shown to be immunogenic entities in contact sensitisation this could be of great use in the development of predictive assays. Peptide sequence(s) could then also be used to identify the intact proteins involved in sensitisation, further aiding the choice of model proteins or design of model peptides for predictive purposes (Divkovic et al., 2005).

References

Ahlfors, S. R., Sterner, O. and Hansson, C. (2003). Reactivity of Contact Allergenic Haptens to Amino Acid Residues in a Model Carrier Peptide, and Characterisation of Formed Peptide-Hapten Adducts. *Skin Pharmacol. Appl. Skin Physiol. 16*, 59-68.

Ahmed, N., Dobler, D., Dean, M. and Thornalley, P. J. (2005). Peptide mapping identifies hotspot site of modification in human serum albumin by methylglyoxal involved in l igand binding and esterase activity. *J. Biol. Chem. 208*, 5724-5732.

Aiba, S., Manome, H., Nakagawa, S. et al. (2003). p38 Mitogen-activated protein kinase and extracellular signal-regulated kinases play distinct roles in the activation of dendritic cells by two representative haptens, NiCl2 and 2,4-dinitrochlorobenzene. *J. Invest. Dermatol. 120*, 390-399.

Alvarez-Sanchez, R., Basketter, D. A., Pease, C, and Lepoittevin, J. P. (2004). Covalent binding of the 13C-labeled skin sensitizers 5-chloro-2-methylisothiazol-3-one (MCI) and 2-methylisothiazol-3-one (MI) to a model peptide and glutathione. *Bioorg. Med. Chem. Lett. 14*, 365-368.

Alvarez-Sanchez, R., Basketter, D. A., Pease, C., and Lepoittevin, J. P. (2003). Studies of chemical selectivity of hapten, reactivity, and skin sensitization potency. 3. Synthesis and studies on the reactivity toward model nucleophiles of the 13C-labeled skin sensitizers, 5-chloro-2-methylisothiazol-3-one (MCI) and 2-methylisothiazol-3-one (MI). *Chem. Res. Toxicol. 16*, 627-636.

Alvarez-Sanchez, R., Divkovic, M., Basketter, D. A. et al. (2004). Effect of Glutathione on the Covalent Binding of the (13)C-Labeled Skin Sensitizer 5-Chloro-2-methylisothiazol-3-one to Human Serum Albumin: Identification of Adducts by Nuclear Magnetic Resonance, Matrix-Assisted Laser Desorption/Ionization Mass Spectrometry, and

Nanoelectrospray Tandem Mass Spectrometry. *Chem. Res. Toxicol. 17*, 1280-1288.

Baker, A., Zidek, L., Wiesler, D. et al. (1998). Reaction of N-acetylglycyllysine methyl ester with 2-alkenals: an alternative model for covalent modification of proteins. *Chem. Res. Toxicol. 11*, 730-740.

Basketter, D. A., Dooms-Goossens, A., Karlberg, A. T. and Lepoittevin, J. P. (1995). The chemistry of contact allergy: why is a molecule allergenic? *Contact Dermatitis 32*, 65-73.

Basketter, D. A., Gerberick, G. F. and Kimber, I. (2001). Measurement of allergenic potency using the local lymph node assay. *Trends Pharmacol. Sci. 22*, 264-265.

Basketter, D. A., Rodford, R., Kimber, I. Et al. (1999). Skin sensitization risk assessment: a comparative evaluation of 3 isothiazolinone biocides. *Contact Dermatitis 40*, 150-154.

Bertucci, C., Ascoli, G., Uccello-Barretta, G. et al. (1995). The binding of 5-fluorouracil to native and modified human serum albumin: UV, CD, and 1H and 19F NMR investigation. *J. Pharm. Biomed. Anal. 13*, 1087-1093.

Bohney, J. P., Fonda, M. L. and Feldhoff, R. C. (1992). Identification of Lys190 as the primary binding site for pyridoxal 5'-phosphate in human serum albumin. *FEBS Lett. 298*, 266-268.

Boisleve, F., Kerdine-Romer, S. and Pallardy, M. (2005). Implication of the MAPK pathways in the maturation of human dendritic cells induced by nickel and TNF-alpha. *Toxicology 206,* 233-244.

Casati, S., Aeby, P., Basketter, D. A. et al. (2005). Dendritic Cells as a Tool for the Predictive Identification of Skin Sensitisation Hazard. *Alternatives to Laboratory Animals 33*, 47-62.

Celis, J. E., Ostergaard, M., Jensen, N. A. et al. (1998). Human and mouse proteomic databases: novel resources in the protein universe. *FEBS Lett. 430*, 64-72.

Collier, P. J., Ramsey, A., Waigh, R. D. et al. (1990). Chemical reactivity of some isothiazolone biocides. *Journal of Applied Bacteriology 69*, 578-584.

Collier, P. J., Ramsey, A. J., Austin, P. and Gilbert, P. (1990). Growth inhibitory and biocidal activity of some isothiazolone biocides. *J. Appl. Bacteriol. 69*, 569-577.

Conduah Birt, J. E., Shuker, D. E. and Farmer, P. B. (1998). Stable acetaldehyde-protein adducts as biomarkers of alcohol exposure. *Chem. Res. Toxicol. 11*, 136-142.

Coombs, R. R. A. and Gell, P. G. H. (1975). Classification of allergic reactions for clinical hypersensitivity and disease. In P. G. H. Gell, R. R. A. Coombs and R. Lachman (eds.), *Clinical Aspects of Immunology* (761-781). London: Blackwell Scientific.

Divkovic, M., Basketter, D. A., Gilmour, N. et al. (2003). Protein-hapten binding: Challenges and limitations for in vitro skin sensitization assays. *Journal of Toxicology – Cutaneous and Ocular Toxicology 22,* 87-99.

Dupuis, G. and Benezra, C. (1982). *Contact Dermatitis to Simple Chemicals: A Molecular Approach*. New York: Marcel Dekker.

Elahi, E. N., Wright Z. M., Hinselwood, D. et al. (2004). Protein Binding and Metabolism Influence the Relative Skin Sensitization Potential of Cinnamic Compounds. *Chem. Res. Toxicol. 17*, 301-310.

Fabrizi, L., Taylor, G .W., Canas, B. et al. (2003). Adduction of the Chloroform Metabolite Phosgene to Lysine Residues of Human Histone H2B. *Chem. Res. Toxicol. 16*, 266-275.

Franot, C., Roberts, D. W., Basketter, D. A. et al. (1994). Structure-activity relationships for contact allergenic potential of gamma,gamma-dimethyl-gamma-butyrolactone derivatives. 2. Quantitative structure-skin sensitization relationships for alpha-substituted-alpha-methyl-gamma,gamma-dimethyl-gamma-butyrolactones. *Chem. Res. Toxicol. 7*, 307-312.

Franot, C., Roberts, D. W., Smith, R. G. et al. (1994). Structure-activity relationships for contact allergenic potential of gamma,gamma-dimethyl-gamma-butyrolactone derivatives. 1. Synthesis and electrophilic reactivity studies of alpha-(omega-substituted-alkyl)-gamma,gamma-dimethyl-gamma-butyrolactones and correlation of skin sensitization potential and cross-sensitization patterns with structure. *Chem. Res. Toxicol. 7*, 297-306.

Gerberick, G. F., Ryan, C. A., Kern, P. S. et al. (2004). A chemical dataset for evaluation of alternative approaches to skin-sensitization testing. *Contact Dermatitis 50*, 274-288.

Girolomoni, G., Gisondi, P., Ottaviani, C. and Cavani, A. (2004). Immunoregulation of allergic contact dermatitis. *J. Dermatol. 31*, 264-270.

Guironnet, G., Dalbiez-Gauthier, C., Rousset, F. et al. (2000). In vitro human T cell sensitization to haptens by monocyte-derived dendritic cells. *Toxicol. In Vitro 14*, 517-522.

Hanson, C., Ahlfors, S. and Sterner, O. (1999). Reactivities of contact allergenic quinones towards a model peptide and identification of the adducts. *Journal of Investigative Dermatology 113*, 320-320.

Hayden, J. B., McCormack, A. L., Yates, J. R., III and Davey, M. P. (1996). Analysis of naturally processed peptides eluted from HLA DRB1*0402 and *0404. *J. Neurosci. Res. 45*, 795-802.

Hopkins, J. E., Naisbitt, D. J., Kitteringham, N. R. et al. (2005). Selective haptenation of cellular or extracellular protein by chemical allergens: association with cytokine polarization. *Chem. Res. Toxicol. 18*, 375-381.

Huang, C.-M., Elmets, C. A., van Kampen, K. R. et al. (2005). Prospective highlights of functional skin proteomics. Mass Spectrom. Rev. e-publication in advance of print.

Krasteva, M., Peguet-Navarro, J., Moulon, C. et al. (1996). In vitro primary sensitization of hapten-specific T cells by cultured human epidermal Langerhans cells – a screening predictive assay for contact sensitizers. *Clin. Exp. Allergy 26*, 563-570.

Landsteiner, K. and Jacobs, J. (1935). Studies on the Sensitisation of Animals With Simple Chemical Compounds. *J. Exp. Med. 61*, 643-656.

Lepoittevin, J. P. (2001). Molecular aspects of allergic contact dermatitis. In R. J. G. Rycroft, T. Menne, P. J. Frosch and J. P. Lepoittevin (eds.), *Textbook of Contact Dermatitis* (59-91). Berlin, Heidelberg: Springer Verlag.

Lippolis, J. D., White, F. M., Marto, J. A. et al. (2002). Analysis of MHC class II antigen processing by quantitation of pep-

tides that constitute nested sets. *J. Immunol. 169*, 5089-5097.

Meschkat, E., Barratt, M. D. and Lepoittevin, J. P. (2001). Studies of the chemical selectivity of hapten, reactivity, and skin sensitization potency. 2. nmr studies of the covalent binding of the (13)c-labeled skin sensitizers 2-[13C]- and 3-[13C]hex-1-ene- and 3-[13C]hexane-1,3-sultones to human serum albumin. *Chem. Res. Toxicol. 14*, 118-126.

Meschkat, E., Barratt, M. D. and Lepoittevin, J. P. (2001). Studies of the chemical selectivity of hapten, reactivity, and skin sensitization potency. 1. Synthesis and studies on the reactivity toward model nucleophiles of the (13)C-labeled skin sensitizers hex-1-ene- and hexane-1,3-sultones. *Chem. Res. Toxicol. 14*, 110-117.

Nerland, D. E., Cai, J. and Benz, F. W. (2003). Selective Covalent Binding of Acrylonitrile to Cys 186 in Rat Liver Carbonic Anhydrase III In Vivo. *Chem. Res. Toxicol. 16*, 583-589.

Nilsson, A. M., Bergstrom, M. A., Luthman, K. et al. (2005). A conjugated diene identified as a prohapten: contact allergenic activity and chemical reactivity of proposed epoxide metabolites. *Chem. Res. Toxicol. 18*, 308-316.

Park, J. H., Lee, Y. ., Kim, K. L. and Cho, E. W. (2003). Selective isolation and identification of HLA-DR-associated naturally processed and presented epitope peptides. *Immunol. Invest 32*, 155-169.

Parker, D., Long, P. V. and Turk, J. L. (1983). A comparison of the conjugation of DNTB and other dinitrobenzenes with free protein radicals and their ability to sensitize or tolerize. *J. Invest Dermatol. 81*, 198-201.

Person, M., Monks, T. J. and Lau, S. S. (2003). An Integrated Approach To Identifying Chemically Induced Post-translational Modifications Using Comparative MALDI-MS and Targeted HPLC-ESI-MS/MS. *Chem. Res. Toxicol. 16*, 598-608.

Peters, T. Jr. (1996). *All about Albumin – Biochemistry, Genetics and Medical Applications*. San Diego: Academic Press.

Pichler, W. J. (2001). Predictive drug allergy testing: an alternative viewpoint. *Toxicology 158*, 31-41.

Rougier, N., Redziniak, G., Mougin, D. et al. (2000). In vitro evaluation of the sensitization potential of weak contact allergens using langerhans-like dendritic cells and autologous T cells. *Toxicology 145*, 73-82.

Rougier, N., Redziniak, G., Schmitt, D. and Vincent, C. (1998). Evaluation of the capacity of dendritic cells derived from cord blood CD34+ precursors to present haptens to unsensitized autologous T cells in vitro. *J. Invest Dermatol. 110*, 348-352.

Rustemeyer, T., von Blomberg, B. M., van Hoogstraten, I. M. et al. (2004). Analysis of effector and regulatory immune reactivity to nickel. *Clin. Exp. Allergy 34*, 1458-1466.

Sharif, S., Mallard, B. A. and Wilkie, B. N. (2003). Characterization of naturally processed and presented peptides associated with bovine major histocompatibility complex (BoLA) class II DR molecules. *Anim Genet. 34*, 116-123.

Smith, C. K. and Hotchkiss, S. A. M. (2001). *Allergic Contact Dermatitis – Chemical and Metabolic Mechanisms*. Taylor & Francis, UK.

Suri, A., Walters, J. J., Kanagawa, O., Gross, M. L. and Unanue, E. R. (2003). Specificity of peptide selection by antigen-presenting cells homozygous or heterozygous for expression of class II MHC molecules: The lack of competition. *Proc. Natl. Acad. Sci. U.S.A. 100*, 5330-5335.

Tracey, B. M. and Shuker, D. E. (1997). Characterization of azo coupling adducts of benzenediazonium ions with aromatic amino acids in peptides and proteins. *Chem. Res. Toxicol. 10*, 1378-1386.

Walker, J. E. (1976) Lysine residue 199 of human serum albumin is modified by acetylsalicyclic acid. *FEBS Lett. 66*, 173-175.

Yasuzawa, T. and Tomer, K. B. (1997). Structural determination of the conjugate of human serum albumin with a mitomycin C derivative, KW-2149, by matrix-assisted laser desorption/ionization mass spectrometry. *Bioconjug. Chem. 8*, 391-399.

Correspondence to
Maja Divkovic
Safety & Environmental Assurance Centre
Unilever Colworth
Sharnbrook
Bedfordshire, MK44 1LQ
UK
e-mail: maja.divkovic@unilever.com

The Chemistry of Skin Allergy

Jean-Pierre Lepoittevin
Université Louis Pasteur Laboratoire de Dermatochimie, Clinique Dermatologique, CHU, Strasbourg, France

Summary
Allergic Contact Dermatitis (ACD) is a very common disease resulting from epidermal proteins being chemically modified by haptens. The processing of such modified proteins by Langerhans cells, the main antigen-presenting cells in the epidermis, generates altered peptides that are subsequently presented, in association with MHC molecules, to naive T-lymphocytes in the lymph nodes. The whole process results in the selection and activation of T-lymphocyte sub-populations with T-cell Receptors (TcR) specific for the chemical modification. Haptens are usually low molecular weight molecules, lipophilic enough to penetrate the epidermis through the stratum corneum, and with a potent chemical reactivity allowing the formation of a covalent link with nucleophilic residues on protein amino acid side chains. For some time it was considered that the more a molecule is able to modify proteins, the better a sensitiser it is, and that a direct relation could be established between the overall chemical reactivity of a molecule and its sensitising potential. Today, it is hypothesised that the sensitising potential of a molecule is related to its chemical reactivity towards a few specific amino acids relevant to the sensitisation process. Haptens can modify proteins by many different mechanisms, from classical nucleophilic-electrophilic reactions to radical reactions. The knowledge of how haptens can modify proteins is the basis for the development of predictive alternative tests aimed at the identification of hazard and potency, such as Structure Activity Relationships (SAR), Quantitative Structure Activity Relationships (QSAR) and peptide reactivity tests.

Keywords: allergic contact dermatitis, haptens, pro-haptens, pre-haptens

Introduction

Allergic contact dermatitis (ACD) is a very common disease resulting from epidermal proteins being chemically modified by haptens (Lepoittevin et al., 1998). The processing of such modified proteins by Langerhans cells, the main antigen-presenting cells in the epidermis, generates altered peptides that are subsequently presented, in association with MHC molecules, to naive T-lymphocytes in the lymph nodes. The whole process results in the selection and activation of T-lymphocyte sub-populations with T-cell receptors (TcR) specific for the chemical modification. Haptens are usually low molecular weight molecules, lipophilic enough to penetrate the epidermis (Roberts and Walter, 1998) through the *stratum corneum* and with a potent chemical reactivity, allowing the formation of a covalent link with nucleophilic residues on protein amino acid side chains. For some time it was considered that the more a molecule can modify proteins, the better a sensitiser it is, and that a direct relation could be established between the overall chemical reactivity of a molecule and its sensitising potential. Today, it is hypothesised that the sensitising potential of a molecule is related to its chemical reactivity towards a few specific amino acids relevant to the sensitisation process. Haptens can modify proteins by many different mechanisms, from classical nucleophilic-electrophilic to radical reactions. The knowledge of how haptens can modify proteins is the basis for the development of predictive alternative tests aimed at the identification of hazard and potency, such as Structure Activity Relationships (SAR), Quantitative Structure Activity Relationships (QSAR) and peptide reactivity tests.

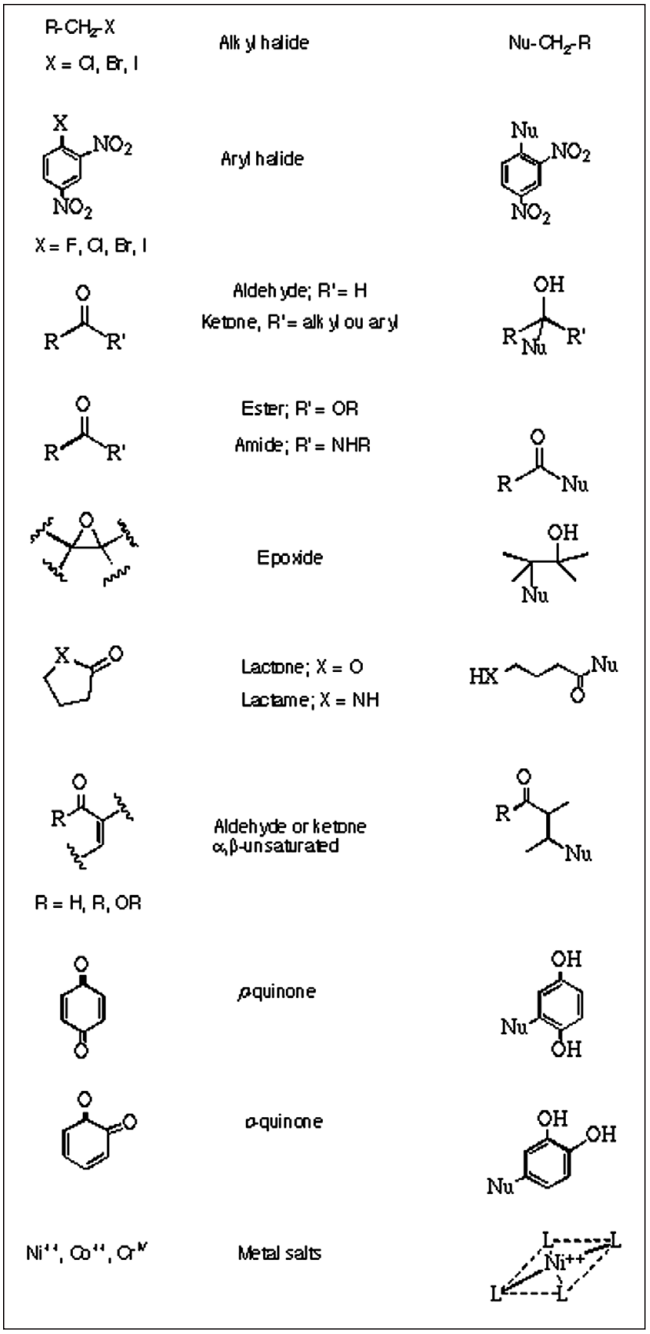
Main electrophilic chemical groups present in contact allergens

Many chemical groups have electrophilic properties and are thus able to react with various nucleophiles to form covalent bonds. Table 1 shows those chemical groups most frequently found in contact allergens.

If we consider biological systems from a chemical viewpoint, it becomes apparent that a very large proportion of structures, especially nucleic acids and proteins, contain many electron-rich groups (those containing nitrogen, phosphorus, oxygen or sulphur). We can thus consider biological systems as being overall nucleophilic. It is therefore not surprising that many biological mechanisms are disturbed on contact with electrophilic chemical substances. Depending on the site of action of these electrophilic molecules, the effect can be mutagenic (Frierson et al., 1985), toxic (Guengerich and Liebler, 1985), or allergenic if the target is the epidermis. In proteins, the side chains of many amino acids contain electron-rich groups capable of reacting with allergens (fig. 2). Lysine and cysteine are those most often cited, but other amino acids containing nucleophilic heteroatoms, such as histidine, methionine, and tyrosine, can react with electrophiles (Means and Feeney, 1971). Thus, it has been shown by nuclear magnetic resonance (NMR) studies that nickel sulphate, for example, interacts with histidine residues of peptides (Romagnoli et al., 1991) and that methyl alkanesulphonates, allergenic methylating agents, mainly react with histidine and to a lesser extent with lysine, methionine, cysteine, and tyrosine (Lepoittevin and Benezra, 1992). If we consider the

chemical structure of some allergens (fig. 2) in the light of the chemical principles already outlined, it is easy to understand that all of these molecules will be able to react with biological nucleophiles. The so formed extremely stable covalent bonds could then lead to the triggering of delayed hypersensitivity.

Tab. 1: Main electrophilic groups seen in contact allergy with the products formed.



Chemical selectivity of haptens for amino acids

A direct consequence of the diversity of hapten-protein interactions is the existence of selectivity for amino acid modifications. For example, we have shown that the α-methylene-γ-butyrolactones, the major allergens of plants of the Asteraceae family, principally modify lysine residues (Franot et al., 1993). It has also been shown that not all modifications were antigenic and that the sensitisation potential of a molecule is probably more related to its ability to modify some specific residues rather than to modify a large number of amino acids. Thus, the difference in

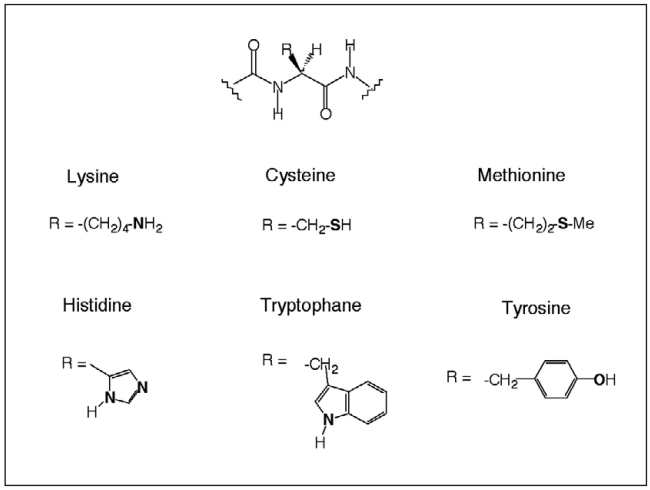


Fig. 1: Principal nucleophilic residues in proteins.

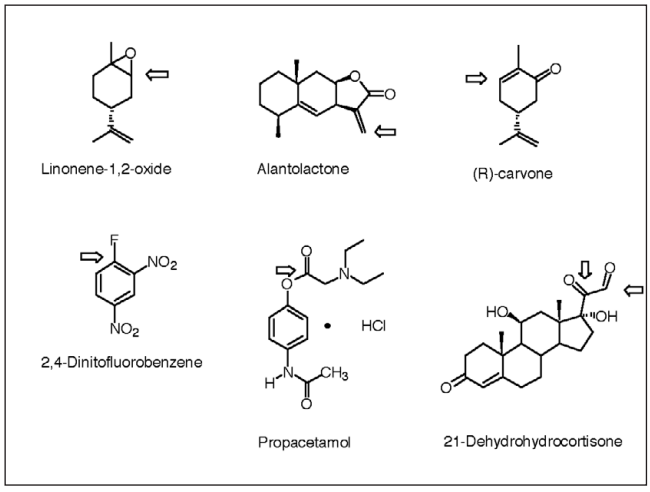


Fig. 2: Examples of sensitising molecules. The electrophilic centre is indicated by an arrow.

sensitising potential of two sultone derivatives, an alkenylsultone (a strong sensitiser) and an alkylsultone (a weak sensitiser), which differ only by the presence of a double bond, could rather be explained by the selective modification of lysine residues by the strong sensitiser than by the many tyrosine residues modified by both derivatives (Meschkat et al., 2001a, 2001b).

A direct consequence of hapten selectivity for amino acids is the existence of differences in the modification sites of proteins and in the density of these modifications. These differences, initially purely chemical, seem increasingly to have a major impact on the response of the immune system. The selectivity of the sites of haptenisation is directly involved in the selection of the peptide fragments that are presented by the APC to the T cells and thus in the selection of T cell receptors. This selectivity also indirectly controls the level of haptenisation of the protein or proteins. It appears that the epitope density on the surface of the APC directly or indirectly directs the immune response towards Th1 or Th2, high epitope densities directing the response towards Th1 and low densities towards Th2 (Hosken et al., 1995). It is reasonable to ask if this selection of response profile, related to epitope density and thus to hapten reactivity, might not explain, for example, the observed differences between respiratory and skin allergens.

In recent years, the radical mechanism has gained increased interest in the discussion of the mechanism of hapten-protein binding. This mechanism, which has never been firmly established, has been postulated to explain, for example, the allergenic potential of eugenol versus iso-eugenol (Barratt, 1992). More recently, studies indicating that radical reactions were important for haptens containing allylic hydroperoxide groups have been published (Lepoittevin and Karlberg, 1994; Giménez-Arnau, 2002).

Metabolism and pro-haptens

Far from being an inert tissue, the skin is the site of many metabolic processes, which can result in structural modification of the xenobiotics that penetrate it (Hotchkiss, 1998). These metabolic processes, primarily intended for the elimination of foreign molecules during detoxification, can, in certain cases, convert harmless molecules into derivatives with electrophilic, and therefore allergenic, properties. The metabolic processes are mainly based on oxidation reactions via extremely powerful enzymatic hydroxylation systems, such as the cytochrome P450 enzymes (Mansuy, 1985), but monoamine oxidases, which convert amines to aldehydes, and peroxidases seem to play an important role in the metabolism of haptens. When activated by the production of hydrogen peroxide during the oxidative stress following the introduction of a xenobiotic into the skin, peroxidases convert the electron-rich aromatic derivatives (aminated or hydroxylated) into quinones, which are powerful electrophiles. In this way, it has been proposed that the long-chain catechols, responsible for the severe allergies to poison ivy (*Rhus radicans* L.) and poison oak (*Rhus diversiloba* T.), could be oxidised *in vivo* to the highly reactive orthoquinones (Dupuis, 1979) (fig. 3). The same applies to paraphenylenediamine or hydroquinone derivatives, such as the allergens from *Phacelia crenulata* Torr. (Reynolds and Rodriguez, 1981), which are converted into electrophilic paraquinones. Metabolic reactions involving enzymatic hydrolyses can also occur in the skin. It is thus that the tuliposides A and B, found in the bulb of the tulip (*Tulipa gesneriana* L.), are hydrolysed, releasing the actual allergens, tulipalines A and B (Bergmann et al., 1967).

All these molecules, which have themselves no electrophilic properties and cannot therefore be haptens but which can be metabolised to haptens, are referred to as pro-haptens (Landsteiner and Jacobs, 1936; Dupuis and Benezra, 1982) and play an important role in contact allergy because of their number and their highly reactive nature. The fact that the structure of the metabolised molecule can be far removed from the structure of the initial molecule can make allergologic investigations even more difficult.

Air oxidation and pre-haptens

Non-enzymatic processes, such as reaction with atmospheric oxygen or ultraviolet irradiation, can also induce changes in the chemical structure of molecules. Many terpenes spontaneously auto-oxidise in air, producing allergising derivatives. In the 1950s it was found that allergenic activity of turpentine was mainly due to hydroperoxides of the monoterpene δ^3-carene (Hellerstrom et al., 1955). This is also the case for abietic acid, the main constituent of colophony, which is converted into the highly reactive substance hydroperoxide (Karlberg, 1988) by contact with air (fig. 4). Such an auto-oxidation mechanism has also been demonstrated for another monoterpene, *d*-limonene, found in citrus fruits. *d*-Limonene itself is not allergenic, but upon exposure to air hydroperoxides, epoxides, and ketones are formed that are strong allergens (Karlberg et al., 1994).

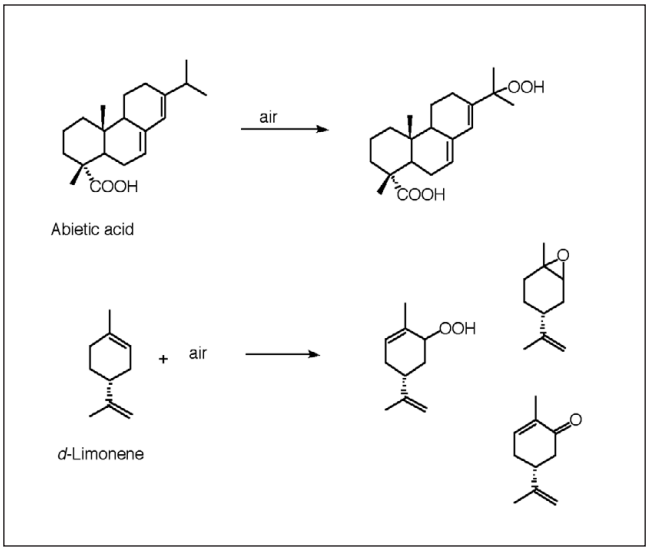


Fig. 3: Examples of pro-hapten metabolism.

All these molecules, which have themselves no electrophilic properties and cannot therefore be haptens but which can be oxidised into haptens, are referred to as pre-haptens.

Conclusion

The principles that we have just discussed permit a rational approach to the phenomena of contact allergy, but, in actual fact, we often have available only indirect evidence suggestive of one mechanism or another. Although the chemical basis for hapten-protein interactions can be checked in the laboratory by the use of nucleophilic amino acids, small peptides and model proteins, and although a certain number of steps can be checked, at the present time no method is available to follow a hapten step by step during the entire immunological process leading to contact allergy. Many points await investigation, but in many cases a "chemical" analysis of the problem does allow us to understand and to foresee cross-allergies and thus to warn the patient about structurally related products.

References

Barratt, M. D. and Basketter, D. A. (1992). Possible origin of the skin sensitization potential of eugenol and related compounds. *Contact Dermatitis 27*, 98-104.

Bergmann, H. H., Beijersbergen, J. C. H., Overeem, J. C. and Sijpesteijn, A. K. (1967). Isolation and identification of α-methylene-γ-butyrolactone: A fungitoxic substance from tulips. *Rec. Trav. Chim. Pays-Bas 86*, 709-713.

Dupuis, G. (1979). Studies of poison ivy. In vitro lymphocytes transformation by urushiol protein conjugates. *Br. J. Dermatol. 101*, 617-624.

Dupuis, G. and Benezra, C. (eds.) (1982). *Allergic contact dermatitis to simple chemicals*. New York: Marcel Dekker.

Franot, C., Benezra, C. and Lepoittevin, J.-P. (1993). Synthesis and interaction studies of 13C labeled lactone derivatives with a model protein using ^{13}C NMR. *Biorg. Med. Chem. 1*, 389-397.

Frierson, M. R., Klopman, G. and Rosenkranz, H. S. (1986). Structure activity relationship among mutagens and carcinogens: A review. *Environ. Mutagen. 8*, 283-327.

Giménez-Arnau, E., Haberkorn, L., Grossi, L. and Lepoittevin, J.-P. (2002). Identification of alkyl radicals derived from an allergenic cyclic tertiary allylic hydroperoxide by combined use of radical trapping and ESR studies. *Tetrahedron 58*, 5535-5545.

Guengerich, F. P. and Liebler, D. C. (1985). Enzymatic activation of chemicals. CRC Crit. *Rev. Toxicol. 14*, 259-307.

Hellerstrom, S., Thyresson, N., Blohm, S.-G. and Widmark, G. (1955). On the nature of eczematogenic component of oxidized delta 3-carene. *J. Invest. Dermatol. 24*, 217-224.

Hosken, N. A., Shibuya, K., Heath, A. W. et al. (1995). The effect of antigen dose on $CD4^+$ T helper cell phenotype development in a T cell receptor-alpha beta-transgenic model. *J. Exp. Med. 182*, 1579-1584.

Hotchkiss, S. A. M. (1998). Dermal metabolism. In M. S. Roberts and K. A. Walters (eds.), *Dermal absorption and toxicity assessment; Drugs and pharmaceutical sciences*, Vol. 91. New York: Marcel Dekker.

Karlberg, A.-T. (1988). Contact allergy to colophony. Chemical identification of allergens. Sensitization experiments and clinical experiments. *Acta Dermato-Venereol. 68 (Suppl. 139)*, 1-43.

Karlberg, A.-T., Shao, L. P., Nilsson, U. et al. (1994). Hydroperoxides in oxidized d-limonene identified as potent contact allergens. *Arch. Dermatol. Res. 286*, 97-103.

Landsteiner, K. and Jacobs, J. L. (1936). Studies on the sensitization of animals with simple chemicals. *J. Exp. Med. 64*, 625-639.

Lepoittevin, J.-P., Basketter, D. A., Goossens, A. and Karlberg A.-T. (eds.) (1998). *Allergic Contact Dermatitis: The molecular basis*, Berlin, Heidelberg, New York: Springer.

Lepoittevin, J.-P. and Benezra, C. (1992). ^{13}C enriched methyl alkanesulfonates: New lipophilic methylating agents for the identification of nucleophilic amino acids of proteins by NMR. *Tetrahedron Lett. 33*, 3875-3878.

Lepoittevin, J.-P. and Karlberg, A.-T. (1994). Interaction of allergenic hydroperoxides with proteins: A radical mechanism? *Chem. Res. Toxicol. 7*, 130-133.

Mansuy, D. (1985). Particular ability of cytochrome P450 to form reactive intermediates and metabolites. In G. Siest (ed.), *Drug metabolism: Molecular approaches and pharmacological implications*. New York: Pergamon.

Means, G. E. and Feeney, R. E. (1971). *Chemical modification of proteins*. San Francisco: Holden Day.

Meschkat, E., Barratt, M. D. and Lepoittevin, J.-P. (2001a). Studies of the chemical selectivity of hapten, reactivity, and skin sensitization potency. 1. Synthesis and studies on the reactivity toward model nucleophiles of the 13C-labeled skin sensitizers hex-1-ene- and hexane-1,3-sultones. *Chem. Res. Toxicol. 14* , 110-117.

Meschkat, E., Barratt, M. D. and Lepoittevin, J.-P. (2001b).

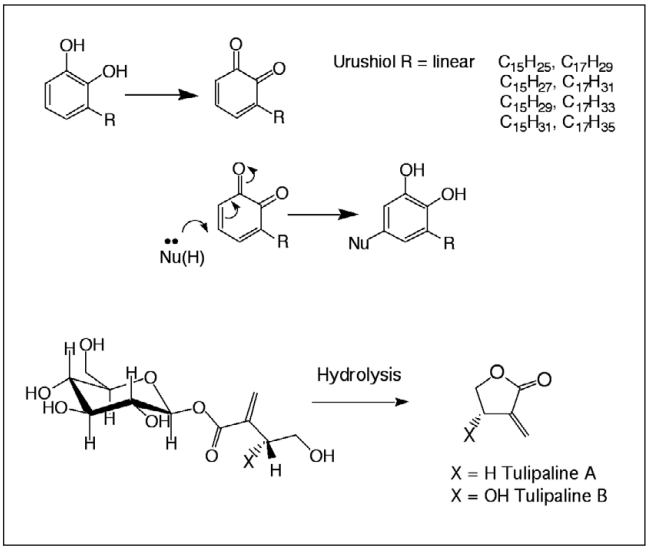


Fig. 4: Examples of chemical modification by reaction with air.

Studies of the chemical selectivity of hapten, reactivity, and skin sensitization potency. 2. NMR studies of the covalent binding of the 13C-labeled skin sensitizers 2-[13C]- and 3-[13C]hex-1-ene- and 3-[13C]hexane-1,3-sultones to human serum albumin. *Chem. Res. Toxicol. 14*, 118-126.

Reynolds, G. and Rodriguez, E. (1981). Prenylated hydroquinones: Contact allergens from trichomes of Phacelia minor and P. parryi. *Phytochemistry 20*, 1365-1366.

Roberts, M. S. and Walters, K. A. (eds.) (1998). *Dermal Absorption and Toxicity Assessment*, New York: Marcel Dekker.

Romagnoli, P., Labahrdt, A. M. and Sinigaglia, F. (1991). Selective interaction of nickel with an MHC bound peptide. *EMBO J. 10*, 1303-1306.

Correspondence to
Prof. Jean-Pierre Lepoittevin
Université Louis Pasteur Laboratoire de Dermatochimie
Clinique Dermatologique, CHU
Strasbourg, France
e-mail: jplepoit@chimie.u-strasbg.fr

Session 5.05
Advancements and needs for developing and validating 3R alternatives for ocular irritancy testing

Use of *In Vitro* Data and (Q)SARs to Classify Eye Irritating Chemicals in the EU – Experience at the BfR

Matthias Herzler, Horst Spielmann, Ingrid Gerner, Manfred Liebsch and Thomas Hoefer
Federal Institute for Risk Assessment (BfR), Berlin, Germany

Summary
Under the EU's proposed new chemicals legislation REACH, the use of experimental animals will be almost completely banned from eye irritation testing of industrial chemicals. Alternative methods are at hand to identify strong irritants. However, there is still a need for reliable tools able to identify non-irritants and to discriminate between non- and moderately irritating chemicals. In addition, testing strategies are required that are capable of integrating all of these approaches with a more efficient exploitation of existing information. In this paper, these issues are addressed, based on regulatory experience gained under the EU New Chemicals notification programme.

Keywords: eye irritation, Draize test, in vitro testing, QSAR, testing strategies

In vivo testing for eye irritation – the Draize test

Today, the Draize rabbit eye test, which was introduced over sixty years ago (Draize et al., 1944), still forms the basis of internationally agreed protocols for eye irritation/corrosion testing (European Commission, 2004; OECD, 2002). The success of this test is based on its obvious biological relevance and the fact that multiple aspects of ocular irritation/corrosion, i.e. different target sites within the eye, as well as the level of severity and the reversibility of effects, are covered in a single test.

However, the Draize test has also been subject to criticism, not only because of the pain and suffering caused to the animals, but also on scientific grounds, among others because of the great variability and low reproducibility of results, an allegedly unrealistic application procedure, and because of differences in the anatomy, physiology, and biochemistry of the rabbit vs. the human eye (Spielmann, 1997; York and Steiling, 1998).

Strategies for classification and labelling

Different systems are in use for the translation of Draize test results data into a classification for eye irritation/corrosion, such as the EU system (European Commission, 2001) or the Globally Harmonised System (GHS, United Nations, 2003). Despite some minor differences (slightly lower thresholds for corneal opacity and conjunctival redness for moderate irritants in the GHS, an additional second sub-category, 2B, for 'mild irritants', is provided for effects that are reversible within 7 days), classification results of these two systems have been proven to be almost identical, while differing significantly from those obtained using other concepts, such as the MMAS approach (Prinsen, 1999).

Today, classification for eye irritation is no longer based on the results of the Draize test alone. Instead, stepwise testing strategies have been developed, such as that summarised in fig-

ure 1, which has been annexed to major internationally accepted test guidelines (European Commission, 2004; OECD, 2002; United Nations, 2003).

In this strategy, the need for a Draize test is waived if any other data or information sources, including physico-chemical considerations, (Q)SARs, and *in vitro* testing, can be used for classification with respect to this hazard. However, in the case of negative or inconclusive results *in silico* or *in vitro*, up to three animals must still be tested. An obvious paradoxon lies in the fact that, according to this approach, most of the Draize tests 'needed' will be carried out with non-irritants which, moreover, are much more common than irritants anyway, at least in the EU New Chemicals database (Hoffmann and Hartung, 2005). Thus, for a large majority of the Draize tests to be performed, a negative test result can be expected.

Demands resulting from the REACH legislation

In 2003, the EU commission published its proposal for a new chemicals legislation, also known as the 'REACH' proposal (European Commission, 2003), which includes among others a fundamental change in the testing strategy for eye irritation that will *de facto* result in an almost total ban of *in vivo* testing of industrial chemicals with respect to this endpoint, if the alternative tools needed to support this new approach can be provided:

- All existing data (*in vitro, in vivo*, historical, validated QSARs, data on chemical analogues) must be considered prior to performing any new test.
- Performing *in vivo* tests will no longer be allowed for obtaining information on corrosive properties (EU risk phrases R34/35), or risk of serious damage to eyes (R41).
- For substances produced or imported in the EU in quantities of 1-10 t/a, an *in vivo* test must not be performed, even for testing potential moderate irritants (EU risk phrase R36). Instead, an animal-free testing strategy will be used, including the evaluation of existing human and animal data, the identification of extreme pH, and a suitable *in vitro* test. Even this test may be waived for known corrosives, if pH is below 2.0 or above 11.5, if the chemical is flammable upon contact with air at room temperature, or if the producer has voluntarily assigned a classification for eye irritation based on skin effects.
- For substances produced or imported in the EU in quantities above 10 t/a, an additional *in vivo* test may only be performed if an inconclusive result was returned by the testing strategy described in the preceding paragraph.

Alternative tools that need to be (further) developed to meet the demands of the future

In order to fulfil the demands of the proposed new policy, the following tools need to be established:

- Simple, animal-free test methods (such as *in vitro* tests or (Q)SARs) for confirming the ABSENCE of irritation/corrosion potential,
- Simple, animal-free test methods (such as *in vitro* tests or (Q)SARs) for discriminating between classification as R41/GHS 1 or R36/GHS 2,
- information management/decision support systems integrating results from different information sources/test methods (*in vivo, in vitro, in silico*) and of different levels of quality to aid industry and regulators in decision-making,
- a comprehensive use of existing information; this would mainly refer to the systematic collection, digitalisation, and evaluation/peer review of already existing toxicological (animal test) data and the development and implementation of efficient data mining techniques.

Regulatory acceptance of such tools will depend on proper validation as well as their ability to translate into internationally accepted classification and labelling systems. However, a number of associated problems arise:

- scoring of eye lesions according to the Draize method is a subjective process resulting in high variability, especially for mild to moderate irritants,
- for validation purposes, these variable Draize scores have to be used, as human reference data are scarce,

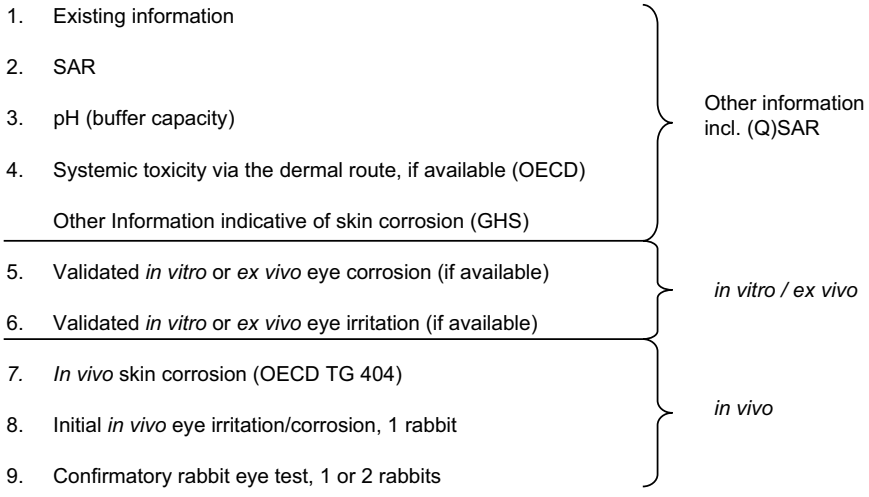


Fig. 1: Current EU/OECD/GHS testing scheme for eye irritation/corrosion

- a principal difference between the Draize rabbit test and alternative techniques lies in the fact that two substances may receive the same classification and labelling based on different effects in different tissues, caused by different mechanisms. In contrast, a single *in vitro* or *in silico* method generally has a more specific focus and will therefore not be able to cover all effects that may lead to a given classification result.

Current acceptance of *in vitro* tests for eye irritation/corrosion in the EU

In the EU, several *in vitro/ex vivo* tests are accepted for the prediction of serious/irreversible eye damage (R41), e.g. (Spielmann 1997):

- the Bovine Corneal Opacity and Permeability (BCOP) test, the Hen Egg Test on the Chorioallantoic Membrane (HET-CAM), the Isolated Rabbit Eye test (IRE), or the Chicken Enucleated Eye Test (CEET).

A survey of data collected at the BfR under the EU New Chemicals notification programme between 1982 and 2004 demonstrates that out of 285 chemicals labelled for risk of serious eye damage (risk phrase R41), 24 had not been tested in the Draize test:

- 10 were classified due to extreme pH (with one also tested in a HET-CAM test),
- 9 were classified according to *in vitro/ex vivo test* results (5 BCOP, 2 IRE, 1 EYETEX, 1 unspecified),
- 3 were classified based on their severe skin irritation potential,
- two more substances were classified based on SAR considerations.

In contrast to serious eye damage, moderate irritation (corresponding to EU risk phrase R36 or GHS Cat. 2) is much more difficult to predict *in vitro*. No validated and generally acknowledged method is available at present. In 1998, at a workshop organised by the European Centre for the Validation of Alternative Methods (ECVAM), experience from previous validation exercises with *in vitro* testing methods for eye irritation/corrosion was evaluated (Balls et al., 1999). As a result, it was recommended to use a battery of tests rather than a single method, to integrate them into step-wise testing strategies, to validate the results against the EU/GHS classification scheme, and – finally – to learn more about the mechanistic basis of eye irritation/corrosion. These conclusions were more or less confirmed in a more recent review (Huggins, 2003).

Use of existing information – the BfR New Chemicals database for acute effects

As described previously (Gerner et al., 2000a; Gerner et al., 2005), quality-reviewed acute systemic and topical toxicity data of about 1,700 substances from the EU New Chemicals notification programme were collected between 1982 and 2002 by Ingrid Gerner at the BfR, primarily as a means of more effectively storing and using existing information for toxicological hazard assessment. All data were assessed for validity and quality by using identical assessment criteria, e.g. only substances of > 95% purity were included, and materials containing impurities suspected to be reactive were excluded. The database was accompanied by a software tool ('Estoff'), for data administration as well as for finding decision rules and structural alerts, and a Decision Support System (DSS) allowing for the application of these rules and alerts to new substances of unknown skin/eye irritation/corrosion potential without providing the actual confidential training set data (Gerner et al., 2000b; Zinke et al., 1999). In detail, the database contains information (if available/applicable) on:

- chemical structure and molecular weight, physico-chemical properties (differentiating between measured vs. estimated or calculated values): pH, aqueous/lipid solubility, melting/boiling points, log P_{OW}, surface tension, vapour pressure,
- potential for hydrolysis and/or thermal decomposition,
- acute systemic toxicity: oral and dermal LD_{50}, inhalative LC_{50}, and EU risk phrases, if assigned,
- skin irritation/corrosion: intensity and reversibility separately for erythemata and oedemata, time to reversibility, and EU risk phrases, if assigned,

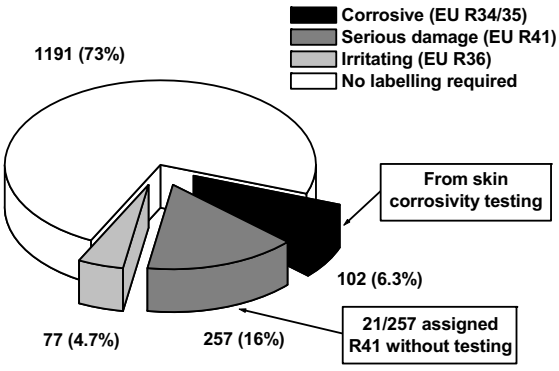


Fig. 2: Distribution of the substances (n = 1627) in the BfR New Chemicals database over different EU classification categories for eye irritation

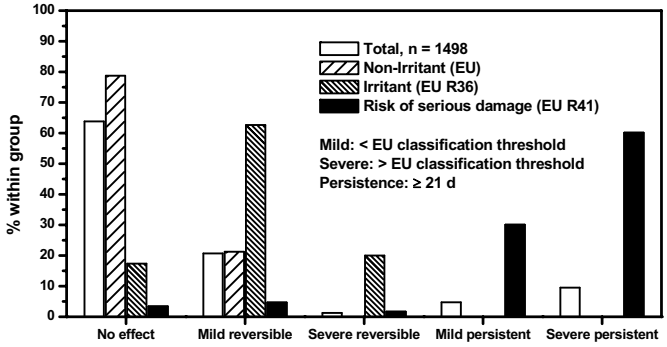


Fig. 3: Discrimination power of the sub-endpoint corneal opacity

- eye irritation/serious damage: intensity and reversibility separately for corneal, iridal, and conjunctival effects, time to reversibility, and EU risk phrases, if assigned,
- sensitisation potential: yes/no, *in vivo* test method used (including percentage of sensitised animals), route (dermal/respiratory).

Figure 2 gives an overview of the distribution of the substances (n=1627) in the database over different EU classification categories for eye irritation.

About 3 out of 4 substances in the database do not require hazard classification with respect to eye irritation/corrosion. This is in good agreement with figures reported by ECVAM from the EU commission's New Chemicals database (Hoffmann and Hartung, 2005). Only about 1 in 5 performed Draize tests actually led to a classification of moderate or severe eye irritation (313 chemicals, i.e. 77 substances classified as R36 and 236 substances classified as R41 based on the Draize test results), while 4 out of 5 tests (1191 substances not classified after a Draize test had been performed) could have been spared, if animal-free test methods/strategies for the exclusion of eye irritation potential had been available.

As already mentioned, the intensity and reversibility of effects are stored separately in the BfR database for the different sub-endpoints covered by the Draize test. These 'existing data' can be used to gain some insight into the potential and limitations of these sub-endpoints for discriminating between different categories of irritants (non-irritant/moderate/severe according to current EU criteria):

Figure 3 demonstrates that corneal opacity is well-suited for discriminating between non- or moderate irritants on the one hand, and severe irritants on the other. A large majority of severe irritants (all R41 substances causing severe reversible or persistent corneal damage, a total of approx. 92%) would have been classified correctly based on corneal opacity scores alone. In contrast, between non-irritants and moderate irritants, there is some degree of overlap and a lower discrimination power can be assumed.

This situation becomes more difficult with conjunctival effects: already almost 2/3 of all non-irritants in the database show some degree of mild, reversible irritation, and the same holds true for substances that were – overall – classified as moderate irritants (fig. 4).

About 80% of all substances and more than 90% of the non-irritants caused at least mild, reversible conjunctival redness (fig. 5).

Finally, for iritis (fig. 6), overlap between all groups is fairly high and, conversely, predictivity for classification becomes rather low, with 20% of the R41 substances displaying either no effect at all or evoking only a sub-threshold reaction. However, in this context it should be noted that iris effects are often camouflaged by the substantial corneal opacity typically caused by this group.

When looking at the data presented in figures 3-6 and considering again that scoring, especially at the border between mild and severe but reversible effects, is subjective and its results are highly variable, it is not surprising that problems with the use of the Draize test (and, consequently, with all alternative methods that are validated against it) must arise when non-irritants have to be discriminated from moderate irritants.

In summary, two conclusions might be drawn from the preceding considerations:

- Difficulties with discriminating between moderate and non-irritants based on Draize test data are an inherent problem of the scoring method and classification thresholds used and not of the alternative methods proposed for replacing the *in vivo* experiment.
- Rather than validating new alternative methods against the overall outcome of a Draize test, it appears more rewarding to perform such validations for each sub-endpoint separately. This approach would also be in line with the previous statement, i.e. that a single *in vitro* test or QSAR model for eye irritation/corrosion will not be able to cover all aspects of a Draize test with the same adequacy. Separate validations could also serve to gain a deeper understanding of the mechanistic aspects of eye irritation/corrosion and to facilitate the identification of the specific potentials and limitations of particular *in vitro* tests.

Using the BfR database to build (Q)SARs

The BfR database was used to establish structural alerts for the prediction of eye irritation potential as well as for building

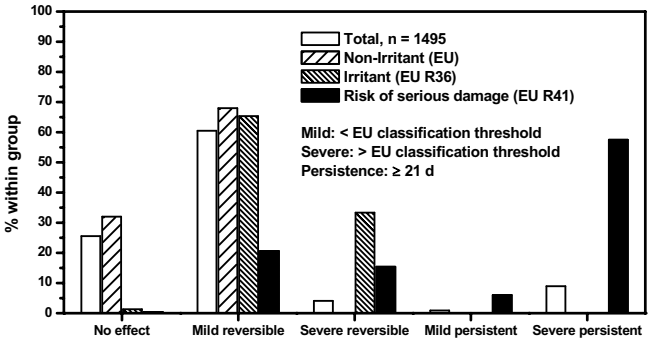


Fig. 4: Discrimination power of the sub-endpoint conjunctival oedema (chemosis)

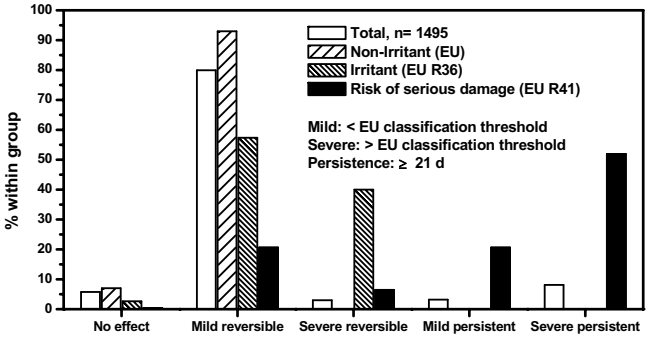


Fig. 5: Discrimination power of the sub-endpoint conjunctival erythema (redness)

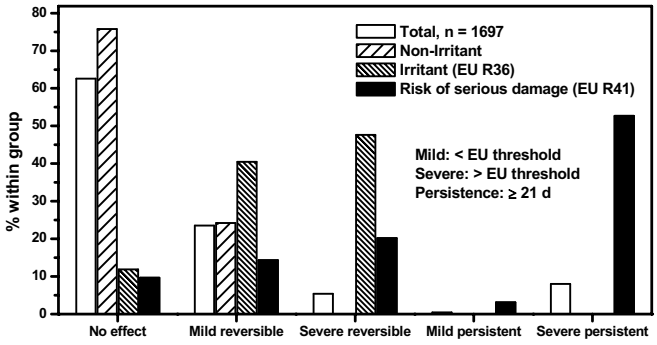


Fig. 6: Discrimination power of the sub-endpoint iritis

physico-chemical rules for the EXCLUSION of such potential. Both alerts and rules have been published before (Gerner et al., 2000b) and were updated recently (Gerner et al., 2005), based on the inclusion of additional data into the database. Examples of some physico-chemical exclusion rules are given in figure 7.

The rules were derived empirically by analysing the distribution of substances over the respective descriptors and setting the limit values such as to exclude 100% of the active (i.e. irritant) substances. Their use is easy and straightforward (in fact, they could be built into a simple Excel spreadsheet): descriptors are not estimated from structural features (e.g. by quantum-mechanical approximations), instead, 'real-life' physico-chemical properties are used for which measured data are readily available, as is at least the case for New Chemicals, pesticides and biocides in the EU.

In 2004, structural alerts and physico-chemical rules for both skin and eye effects have been submitted for validation to the (Q)SAR group at the European Chemicals Bureau (ECB). This validation project might be seen as a model for future validation exercises with models built on confidential data: The collection of data and model development were performed in one EU member state (BfR, Germany), the process was organised by a supranational European body (EU Joint Research Centre/ECB), and the validation itself was contracted to independent experts from a second EU member state (RIVM, the Netherlands).

In the meantime, the first part of the validation, i.e. the validation of skin irritation/corrosion physico-chemical rules has been completed (Rorije and Hulzebos, 2005), and some preliminary conclusions can be drawn that will most likely also apply to the rules for eye irritation:

- In general, the rules are in good agreement with the OECD principles for (Q)SAR validation.
- High predictivity was obtained in an external validation with a second dataset of EU New Chemicals (also compiled at the BfR, but not used to build the rules).
- Skin irritation/corrosion testing for up to approx. 40% of the test set substances could have been waived based on the validated rules. However, this percentage might be lower for test datasets from other domains of the chemical space (e. g. pharmaceuticals, pesticides, etc.).

After successful validation, the exclusion rules are proposed to act as early-stage 'exclusion filters' that are combined with both structural alerts and *in vitro* testing under the frame of an integrated testing strategy, such as that depicted in figure 8 (Gerner et al., 2005; Gerner and Schlede, 2002).

Recommendations for future work

Improved *in vitro* tools need to be developed for classification of moderate eye irritation (EU R36/GHS Category 2) as well as for predicting the absence of irritation potential. Internationally recognised test guidelines and standard operating procedures are required for the harmonised use of *in vitro* tests to predict eye irritation/corrosion. Validation of both the structural alerts and physico-chemical exclusion rules developed at the BfR should be extended to substances outside of the 'New Chemicals Space', e.g. pesticides, biocides or cosmetic ingredients. Further exploitation of the BfR database with different

Rules appropriate for all groups of chemicals

Basis
Evaluation of data for 1627 chemicals with purity ≥ 95%

If melting point > 200 ºC	Then not (skin corrosion R34 or R35) (true for 245/252 chemicals tested)[a]
If log P_{ow} > 9	Then not (lesions R34, R35, R36 or R41) (true for 32/32 chemicals tested)
If log P_{ow} < -3.1	Then not (skin corrosion R34 or R35) (true for 53/53 chemicals tested)
If lipid solubility < 0.01 g/kg	Then not (skin corrosion R34 or R35) (true for 58/58 chemicals tested)
If aqueous solubility < 0.00002 g/l	Then not (eye irritation R41) (true for 109/109 chemicals tested)
If aqueous solubility < 0.000005 g/l	Then not (eye irritation R36) (true for 38/38 chemicals tested)
If molecular mass > 650 g/Mol	Then not (eye irritation R36) (true for 139/139 chemicals tested)[b]

[a] *The seven skin corrosive substances are organic salts which release strong inorganic acids or bases when in contact with aqueous substrates/organic media*

[b] *Chemicals with molecular mass > 650 g/Mol may elicit severe tissue damage resulting in local corrosion (labelled R41)*

Fig. 7: Example of physico-chemical exclusion rules derived from the BfR database (from Gerner et al., 2005, where a more detailed description can be found)

(multivariate) statistical approaches could give deeper insight into the interdependency of both the physico-chemical descriptor variables and the different sub-endpoints of the Draize test (cf. e.g. Lovell, 1996). More sophisticated concepts to exploit existing *in vivo* test data more efficiently should be developed. A large collection of historical animal test data is stored in the archives of both industry and governmental authorities – a treasure of information still waiting to be systematically evaluated for its potential to be used as an aid to further reduce animal testing for eye irritation.

Summary and conclusions

In order to avoid the use of experimental animals in eye irritation/corrosion testing, integrated testing strategies are needed

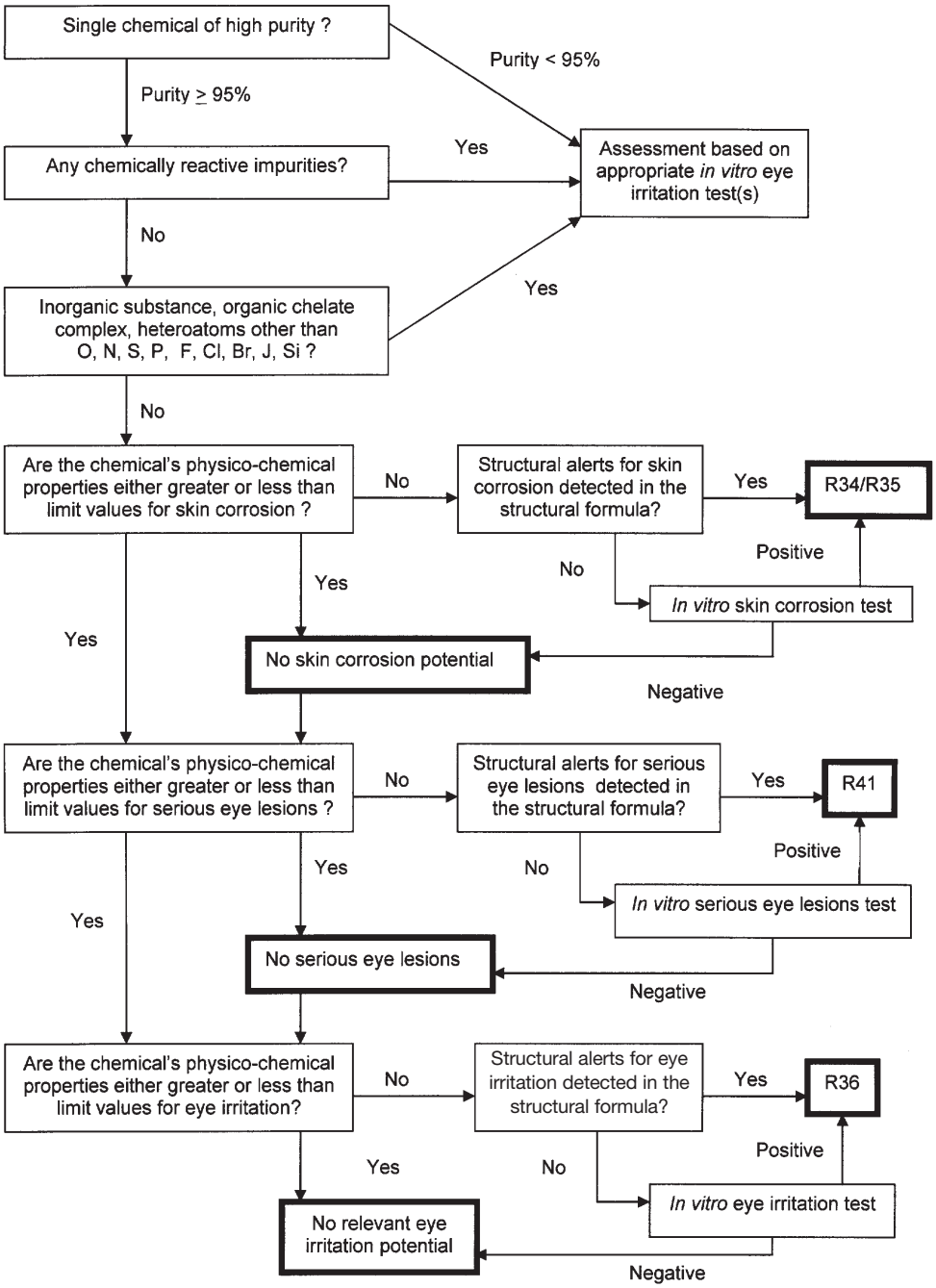


Fig. 8: Proposed integration of physico-chemical exclusion rules, structural alerts and *in vitro* testing in an animal-free testing strategy for skin/eye irritation/corrosion

that combine the use of existing information, chemical knowledge, batteries of *in vitro* tests and (Q)SARs. For the prediction of serious/irreversible eye damage, validated *in vitro* methods are at hand that may most efficiently be used in a battery approach, however, valid methods are missing for reliable identification of moderate and non-irritants. During 20 years of New Chemicals notification in the EU, a database on the acute systemic and topical toxicity of some 1,700 chemicals was compiled at the BfR. From this database, structural alerts for the prediction of eye irritation/corrosion potential were derived as well as physico-chemical exclusion rules to predict the ABSENCE of such effects. Both the alerts and exclusion rules were submitted to the ECB for validation and first experience obtained with the validation of exclusion rules for skin irritation/corrosion produced promising results.

References

Balls, M., Berg, N., Bruner, L. H. et al. (1999). Eye irritation testing: the way forward. *Alternatives to Laboratory Animals 27*, 53-77.

Draize, J. H., Woodard, G. and Calvery, H. O. (1944). Methods for the study of irritation and toxicity of substances applied topically to the skin and mucous membranes. *Journal of Pharmacology and Experimental Therapeutics 82*, 377-390.

European Commission (2001). Commission Directive 2001/59/EC of 6 August 2001 adapting to technical progress for the 28th time Council Directive 67/548/EEC on the approximation of the laws, regulations and administrative provisions relating to the classification, packaging and labelling of dangerous substances, OJ L225/1.

European Commission (2003). Proposal for a regulation of the regulation of the European Parliament and of the Council concerning the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH), establishing a European Chemicals Agency and amending Directive 1999/45/EC and Regulation (EC) {on Persistent Organic Pollutants}, COM (2003) 644 final.

European Commission (2004). Commission Directive 2004/73/EC of 29 April 2004 adapting to technical progress for the 29th time Council Directive 67/548/EEC on the approximation of the laws, regulations and administrative provisions relating to the classification, packaging and labelling of dangerous substances, OJ L216/3.

Gerner, I., Graetschel, G., Kahl, J. and Schlede, E. (2000a). Development of a decision support system for the introduction of alternative methods into local irritancy/corrosivity testing strategies. Development of a relational database. *Alternatives to Laboratory Animals 2*8, 11-28.

Gerner, I., Liebsch, M. and Spielmann, H. (2005). Assessment of the eye irritating properties of chemicals by applying alternatives to the Draize rabbit eye test: the use of QSARs and in vitro tests for the classification of eye irritation. *Alternatives to Laboratory Animals 33*, 215-237.

Gerner, I. and Schlede, E. (2002). Introduction of in vitro data into local irritation/corrosion testing strategies by means of SAR considerations: assessment of chemicals. *Toxicology Letters 127*, 169-175.

Gerner, I., Zinke, S., Graetschel, G., and Schlede, E. (2000b). Development of a decision support system for the introduction of alternative methods into local irritancy/corrosivity testing strategies. Creation of fundamental rules for a decision support system. *Alternatives to Laboratory Animals 28*, 665-698.

Hoffmann, S. and Hartung, T. (2005). Prevalence and test interdependence: pivotal parameters in the design and validation of testing strategies. *ALTEX 22*, 281.

Huggins, J. (2003). Alternatives to animal testing: research, trends, validation, regulatory acceptance. *ALTEX Suppl. 1/03*, 3-61.

Lovell, D. P. (1996). Principal component analysis of Draize eye irritation tissue scores from 72 samples of 55 chemicals in the ECETOC data bank. *Toxicology in Vitro 10*, 609-618.

OECD (2002). Test guideline 405 "Acute Eye Irritation/ Corrosion".

Prinsen, M. K. (1999). An Evaluation of the OECD Proposal for the Harmonised Classification of Eye Irritants and Corrosives. *Alternatives to Laboratory Animals 27*, 72-77.

Rorije, E. and Hulzebos, E. M. (2005). Evaluation of (Q)SARs for the Prediction of Skin Irritation/Corrosion Potential. Unpublished report. National Institute of Public Health and Environment, Expert Centre for Subsatnces (RIVM-SEC). Bilthoven, The Netherlands.

Spielmann, H. (1997). Ocular Irritation. In J. V. Castell and M. J. Gómez-Lechón (eds.), *In vitro methods in pharmaceutical research* (265-287). San Diego/CA, USA: Academic Press.

United Nations Economic Commission for Europe/UNECE (2003). Globally Harmonized System of Classification and Labelling of Chemicals (GHS), 137-149.

York, M. and Steiling, W. (1998). A critical review of the assessment of eye irritation potential using the Draize rabbit eye test. *Journal of Applied Toxicology 18*, 233-240.

Zinke, S., Gerner, I., Graetschel, G., and Schlede, E. (1999). Local irritation/corrosion testing strategies: development of a decision support system for the introduction of alternative methods. *Alternatives to Laboratory Animals 28*, 29-40.

Correspondence to

Matthias Herzler
Federal Institute for Risk Assessment (BfR)
Thielallee 88-92
14195 Berlin
Germany
e-mail: m.herzler@bfr.bund.de

ICCVAM Progress in Evaluating *In Vitro* Test Methods for Identifying Severe Ocular Irritants/Corrosives

Jill Merrill[1], Karen Hamernik[2], Leonard Schechtman[1], William Stokes[3] and Marilyn Wind[4]

[1]US Food and Drug Administration, Center for Drug Evaluation and Research, Rockville, MD, USA; [2]U. S. Environmental Protection Agency, Washington, DC, USA; [3]National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods, National Institute of Environmental Health Sciences, NIH, DHHS, Research Triangle Park, NC, USA; [4]Consumer Product Safety Commission, Bethesda, MD, USA

The views expressed in this paper do not necessarily represent the official positions of any US Federal government agencies.

Summary

In response to a nomination by the US Environmental Protection Agency, the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) initiated a review of the validation status of four in vitro test methods for use in screening chemicals for severe eye irritation or corrosion. The four test methods are the isolated rabbit eye test (IRE), isolated chicken eye test (ICE), hen's egg test on chorioallantoic membrane (HET-CAM) and bovine corneal opacity and permeability assay (BCOP). Background review documents (BRDs) were prepared based on the available data and independently reviewed by an Expert Panel. This paper describes the status of the review process as of August, 2005.

Keywords: bovine corneal opacity and permeability assay (BCOP), hen's egg test on chorioallantoic membrane (HET-CAM), ICCVAM, in vitro ocular toxicity test methods, isolated chicken eye test (ICE), isolated rabbit eye test (IRE)

Introduction

An ongoing ICCVAM review is being conducted with the overall goal of evaluating the validation status of four *in vitro* test methods (IRE, ICE, HET-CAM and BCOP) for their possible regulatory use in a tiered-testing strategy. The use of these tests, once appropriately validated, could reduce and refine animal use. Using a tiered-testing strategy, substances that tested positive in a validated *in vitro* test could be classified and labelled as severe ocular irritants/corrosives with no *in vivo* testing necessary. Substances that tested negative would undergo additional testing, either in the *in vivo* rabbit eye test or another validated *in vitro* test capable of detecting false negatives in the first *in vitro* test (This testing strategy may not apply to some pharmaceuticals). However, the key to this tiered-testing strategy is the availability of appropriately validated *in vitro* tests.

Validation of alternative methods in the United States

In the United States, ICCVAM, which is composed of 15 federal agencies, coordinates the technical review of new or revised alternative test methods as well as issues related to their validation. An important aim of the ICCVAM process is to help facilitate regulatory acceptance of such methods, as appropriate, by relevant member agencies. Priority for ICCVAM activities is generally given to methods that may improve the prediction of adverse human, animal or ecological effects and those that might reduce, refine or replace animal use. Validation has been defined by ICCVAM as the process by which the reliability and relevance of an assay for a specific purpose is established (ICCVAM, 1997). Reliability is defined as the reproducibility of the test method within and among different laboratories. It should be based on performance with different substances, representing the types of chemicals and product classes that are expected to be tested, and the range of responses that needs to be identified. Relevance is defined as the extent to which an assay will correctly predict or measure the biological effect of interest (i.e., via performance characteristic measures compared to a standard). (In the draft version of the BRDs the term "accuracy" rather than the term "relevance" was used to describe the overall performance characteristics of a test method as the comparison was made with respect to rabbit *in vivo* data). Ideally this analysis would have evaluated the ability of the *in vitro* test to correctly predict human ocular toxicity testing. However, severe ocular irritants are not tested in human eyes and in the absence of such data, how well the *in vitro* test predicts the *in vivo* rabbit test was measured.

A technical review was initiated to assess the current validation status of each of the four test methods. An ICCVAM Ocular Toxicity Working Group (OTWG) was established to work with the NTP (National Toxicology Program) Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM) to carry out these evaluations. NICEATM provides support to ICCVAM. ICCVAM also collaborated closely with the European Centre for the Validation of Alternative Methods

(ECVAM). Draft BRDs were prepared for each of the four test methods. The BRDs were designed to be comprehensive, reviewing the available data and information for each of the four methods. As such, the information included in the BRDs describes what is known about the accuracy and reliability of the test method, the scope of substances tested and the availability of a standardised protocol. In addition, based on existing data, the BRDs describe the known usefulness and limitations of each test method. Recommendations for any future optimisation or validation studies and a list of recommended reference substances for those validation studies as well as a standardised test method protocol are included in each BRD.

In order to assess the validation status of the test methods, high quality *in vivo* rabbit eye data and high quality *in vitro* data for each test method was needed. Some of this came from the peer-reviewed literature and some was submitted to NICEATM in response to a Federal Register (FR) notice (March 24, 2004). To be considered in the analysis, the *in vivo* data had to be based on the Draize scoring system and had to meet the following acceptance criteria: At least three rabbits were tested in the study unless a severe effect was noted in a single animal, in which case substance classification could be based on effects observed in less than three animals; 0.1 mL or 0.1 g was tested in each animal, unless a severe effect was noted with lesser amounts; minimally, observations had to have been made at 24, 48, and 72 h, unless a severe effect was observed earlier. If any of the above three criteria were not met, the data for that substance could not be used for the accuracy analyses.

The goal was to be able to classify each substance using the three major hazard classification systems: the US Environmental Protection Agency (EPA), the European Union (EU) and the UN Globally Harmonized System (GHS). However, each of these systems uses different decision criteria to identify corrosives/severe irritants based on the same rabbit data. For example, although all three systems make classifications based on the magnitude of the individual rabbit response, the EPA and the GHS also use the time taken for ocular lesions to clear. Therefore, individual rabbit data, collected at the different observation times, were needed if the *in vivo* data were to be used in the accuracy analyses. This was not always available, resulting in some submitted data not being included in the accuracy analyses. The accuracy of each proposed test method, compared to the *in vivo* rabbit test, was evaluated by calculating various statistics, of which the key ones are described below. The first of these is also termed accuracy, but in this context it refers to accuracy in the sense of concordance and measures the proportion of correct outcomes, either positive or negative, of a test method. Sensitivity measures the proportion of all positive substances that are correctly classified as positive relative to the standard of comparison, whereas specificity measures the proportion of all negative substances that are correctly classified as negative. The false positive rate measures the proportion of all negative substances that are incorrectly identified as positive, whereas the false negative rate measures the proportion of all positive substances that are incorrectly identified as negative. Reliability includes intra-laboratory repeatability and intra- and inter-laboratory reproducibility. Repeatability refers to the closeness of agreement between results on the same substance, using identical conditions within a given time period. Intra-laboratory reproducibility refers to the extent to which qualified personnel, within the same laboratory, can replicate results using a specific protocol at different time periods. Inter-laboratory reproducibility refers to the extent to which qualified personnel can replicate results in different laboratories (i.e., transferability).

Based on each of the three classification systems, an accuracy and reliability analysis was performed for each of the four test methods, except where the available data did not permit a complete analysis. Once these analyses were complete, the draft BRDs were made publicly available on the ICCVAM/NICEATM website (http://iccvam.niehs.gov) in November 2004.

Expert Panel review of the BRDs

The BRDs were then reviewed by an independent international expert panel. The Expert Panel was selected with input from ICCVAM, the OTWG and ECVAM. Members of the Panel are recognised experts in the field and come from different backgrounds with representatives from academia, government, industry and animal welfare. Panel members were specifically chosen for both their scientific expertise and their lack of direct involvement with the test methods under consideration. A public meeting of the Panel was held on January 11-12, 2005 at the National Institutes of Health, Bethesda, Maryland. The Panel was charged with evaluating the extent and adequacy that each of the applicable ICCVAM validation and acceptance criteria had been addressed, based on the available information and data, or will be addressed in proposed studies. They were also charged with developing conclusions and recommendations on the usefulness and limitations of the assays, the protocol that should be used for any future testing and validation studies, the adequacy of the proposed optimisation and/or validation studies and the adequacy of reference substances proposed for future validation studies.

Results of the Expert Panel review

Although the Panel's final report is available on the ICCVAM/NICEATM website (http://iccvam.niehs.nih.gov), the key points for each test method will be discussed here.

IRE

Based on the available data, an evaluation of intra-lab repeatability and reproducibility could not be performed for this method. The Panel concluded that the BRD-proposed test method appears useful in a tiered-testing strategy to identify severe ocular irritants/corrosives. However, the test method accuracy must be corroborated with a larger number of substances and a reliability analysis should be conducted when additional data become available. With respect to optimisation and validation, the Panel recommended additional data be

requested from the test method users and that a reanalysis of all the data be conducted subsequently. The Panel also proposed several modifications to the standardised protocol, including the use of positive and negative controls and benchmark substances, the inclusion of methods to detect cellular damage and death, the development of a standardised histopathology scoring system for corneal damage, and the use of reference photographs for all subjective endpoints.

ICE

The Panel concluded the ICE test method can be used in a tiered-testing strategy to identify severe ocular irritants/corrosives. However, the Panel noted that alcohols tend to be overpredicted, while surfactants tend to be underpredicted. The Panel suggested that solids and insoluble substances may not come in adequate contact with the corneal surface, resulting in underprediction. The Panel also proposed several modifications to the proposed protocol, including adding centering lights to the optical pachymeter; the use of histopathology when the standard ICE endpoints produce borderline results; the use of reference photographs for all subjective endpoints; the use of concurrent negative and positive control eyes (at least three eyes per group). The Panel suggested control eyes be spread throughout the superperfusion apparatus such that replicate eyes are placed randomly. This would make order effects in dosing less likely. Given the limited amount of ICE reliability data, additional studies were suggested to better characterise the repeatability and reproducibility of the test method.

HET-CAM

The Panel concluded that HET-CAM is useful in a tiered-testing strategy to identify severe ocular irritants/corrosives. However, the high false positive rate is a limitation of the test method. They suggested retesting positive HET-CAM results in a modified HET-CAM or in a different *in vitro* test method. The Panel stated that optimisation studies could increase the accuracy of HET-CAM, possibly reducing the false positive rate while maintaining an acceptable false negative rate. Therefore, a retrospective analysis should be conducted to determine whether different decision criteria might enhance the accuracy and/or reliability of the test method. The Panel also proposed modifications to the proposed protocol, including the inclusion of different endpoints (e.g. trypan blue absorption, antibody staining, membrane changes, etc.) which may reduce the false positive rate; the inclusion of procedures for applying and removing solids from the chorioallantoic membrane (CAM) which otherwise may adhere to the CAM and damage it upon removal.

BCOP

The Panel concluded that the BRD-proposed test method is useful in a tiered-testing strategy for identifying severe ocular irritants/corrosives. However, the test should not be used for alcohols, ketones, and solids. Further optimisation and validation studies are necessary before these materials can potentially be assessed with this assay. It needs to be confirmed that the BCOP identifies known human ocular irritants as well as or better than the Draize test. The Panel concluded that histology should be added to the test method protocol, unless the test substance is from a class of materials known to be accurately predicted using only opacity and permeability. The Panel also expressed concern that users be aware of the possibility of zoonoses, including Bovine Spongiform Encephalopathy. They recommended that standard universal precautions, such as gloves and glasses, always be used. The Panel also proposed several modifications to the proposed protocol including the use of the holders suggested by Ubels et al., (2002); re-examining the use of the calculated total score when the endpoint is severe injury only; possible changes to the medium used to bathe the eyes, including determination of whether foetal bovine serum is needed. The Panel also suggested the possibility of using the porcine eye as a model for the human eye. They recognised that this change would require a complete validation, but wanted to be sure that it was considered for future work.

For all four methods, the Panel specified that any further optimisation or validation studies should be conducted using existing animal data. Additional animal studies should only be conducted if important data gaps are identified. Such studies would need to be carefully designed to maximise the amount of information obtained and minimise animal usage. The Panel also commented on the accuracy and reliability of the *in vivo* rabbit test, stating there should be more discussion of the variability of the *in vivo* rabbit data. This is particularly important in the determination of accuracy of an *in vitro* test method. Because of the known variability in the rabbit test, it is not possible from the data presented in the BRD to determine whether the inconsistencies between the two tests are due to 'failure' of the *in vitro* test method or a misclassification by the single *in vivo* test result. Some public comments also expressed concern that the variability of the rabbit data be more prominently discussed in the BRDs.

The Panel also reviewed the adequacy and completeness of the proposed list of reference substances. They concluded that the list is comprehensive, substances appear to be commercially available in an acceptably pure form, and an appropriate range of ocular responses appears adequately represented. Although the Panel recognised that the list is limited by the availability of *in vivo* reference data, they concluded that surfactants are overrepresented, the list is too long and should be shortened, but at the same time more inorganic substances should be added. They also recommended that substances known to induce severe ocular lesions in humans be included in the list, even in the absence of rabbit data.

Additional data and data reanalysis

Public comments made at the meeting indicated that additional data could be made available. The Expert Panel recommended that this data be requested and that a reanalysis of the accuracy and reliability of each test method be conducted. Consequently a second FR notice was published February 28, 2005, requesting all available *in vitro* data on these test methods and any corresponding *in vivo* rabbit data.

An accuracy reanalysis was also required because, after the

BRDs were released, NICEATM received clarification of the classification rules for severe irritants. This change resulted in a small number of substances previously classified as non-severe now being classified as severe. A reanalysis was also required because a standardised chemical structure classification scheme, based on Medical Subject Headings, was used to ensure consistency in classifying substances tested in the *in vitro* ocular test methods. This resulted in some chemicals being reclassified.

Revised performance characteristics

A brief summary of the revised performance characteristics is provided for each test method in table 1. These characteristics are important in considering a method's usefulness in a tiered-testing strategy. This summary is based on the GHS classification system only because the international community appears to be adopting this classification system. Based on the reanalysis, IRE had a sensitivity of 100%, identifying all known corrosives as corrosive. However, it was overly sensitive, with a high false positive rate (56%). Based on the limited number of substances evaluated, IRE had a false negative rate of 0% and may be useful in identifying a substance as non-corrosive. ICE had a low false positive rate and as such it appears to have utility in reliably identifying a substance as corrosive. But, with a sensitivity of 50%, it missed many corrosives. The high false negative rate means that products sent on for confirmatory *in vivo* testing have a high possibility of being severely irritating to rabbit eyes. HET-CAM had a high false positive rate (60%). However, the assay may still be useful in a tiered-testing strategy where positive substances could be retested *in vivo* or in another appropriately validated alternative method to confirm the result. The BCOP assay was fairly good at predicting both known corrosives and non-corrosives. It was also fairly reliable at correctly identifying a product as corrosive with a false positive rate of 20%. In terms of identifying a product as non-corrosive it was fairly reliable with a false negative rate of 16%.

The results of the reanalyses were made publicly available on the ICCVAM/NICEATM website (http://iccvam.niehs.nih.gov) in July, 2005. The Expert Panel has since been asked whether the results of the reanalyses form the basis for any changes to their conclusions. Once the Panel's conclusions are final, ICCVAM and OTWG will consider their report and the public comments received in response to the review process. ICCVAM will then make recommendations regarding these test methods to US Federal Agencies for their consideration and action.

References

ICCVAM (1997). Validation and regulatory acceptance of toxicological test methods: A Report of the ad hoc Interagency Coordinating Committee on the Validation of Alternative Methods. NIH Publication No.: 97-3981. Research Triangle Park, NC: National Institute of Environmental Health Sciences.

Ubels, J. L., Paauw, J. D., Casterton, P. L. and Kool, D. J. (2002). A redesigned corneal holder for the bovine cornea opacity and permeability assay that maintains normal corneal morphology. *Toxicol. In Vitro 16*, 621-628.

Acknowledgements

This review represents a collaborative effort between ICCVAM, ECVAM, OTWG and NICEATM. Many people participated with public comments, providing useful information and constructive criticism. The authors would like to acknowledge the time and expertise of all these people and thank them for their hard work in this ongoing collaborative effort.

Correspondence to

Jill Merrill, PhD, DABT
US Food and Drug Administration
Center for Drug Evaluation and Research
Rockville, MD, USA
e-mail: jill.merrill@fda.hhs.gov

Session 5.06
In vitro approaches for determining acute systemic toxicity

Estimating Acute Toxicity Based on *In Vitro* Cytotoxicity: Role of Biokinetic Modelling

Bas J. Blaauboer[1], Joop Hermens[1] and Jan van Eijkeren[2]
[1]Institute for Risk Assessment Sciences, Utrecht University, The Netherlands; [2]National Institute of Health and the Environment (RIVM), Bilthoven, The Netherlands

Summary
Deviations from a linear relationship between cytotoxic concentrations of a substance in vitro and toxic doses in vivo can result from the fact that effective concentrations in vitro are irrelevant for the concentrations that cause toxicity in target organs in vivo. An important reason is a compound's biokinetic behaviour, which determines the concentrations reached in target organs.
Physiologically-based biokinetic modelling will be employed to alert to chemicals displaying a deviating kinetic behaviour and that will thus need further attention to estimate their acute toxicity. This will improve the accuracy of estimating the in vivo toxic dose on the basis of in vitro basal cytotoxicity.

Keywords: acute toxicity, biokinetic modelling, cytotoxicity

Introduction

In the search for methods to evaluate the toxicological risk of chemicals without employing animal experimentation, much emphasis has been put on the replacement of acute toxicity (LD_{50}) determinations. One important assumption was that acute toxicity is related to a compound's basal cytotoxicity (Ekwall, 1983). In previous programmes, the feasibility of the use of *in vitro* cytotoxicity data for the prediction of *in vivo* lethal doses was tested, e.g. in the MEIC study (Ekwall et al., 1998), and by publication of the Register of Cytotoxicity (Spielmann et al., 1999; Halle., 2003). In these studies it was shown that basal cytotoxicity data resulted in good estimates for about 70% of the compounds, i.e. these chemicals could be classified in the appropriate LD_{50} classes (Clemedson and Ekwall, 1999). This implies that for about 30% of the cases the estimates on the basis of cytotoxicity data deviated from *in vivo* findings.

These deviations from a simple linear relationship between effective concentrations *in vitro* and toxic doses *in vivo* can result from the fact that the effective concentrations *in vitro* are irrelevant for the concentrations that may cause toxicity at the target site in target organs *in vivo*. This will in many cases be the result of the biokinetics of the compound under study (Blaauboer, 2002). For instance, the absorption of the compound may be minimal, thus leading to low systemic concentrations. Moreover, the processes of distribution, metabolism and elimination may lead to lower or higher concentrations in target organs than could be expected from an even distribution of the compound in the body. Thus, an important drawback of the use of cytotoxicity data is the difficulty of extrapolating a toxic concentration in the *in vitro* system to a toxic dose in the *in vivo* situation (Blaauboer, 2003).

The ACuteTox programme aims to improve the estimates for acute toxicity on the basis of non-animal studies (e.g. *in vitro* cytotoxicity studies). Thus, one of the areas needing attention is the study of the biokinetics of the test compounds. Two work packages in the programme are devoted to this: work package 5 involves studies of absorption, distribution and elimination, while work package 6 is devoted to the role of biotransforma-

tion. Other parts of the programme are devoted to other possible reasons for the absence of a clear relationship between basal cytotoxicity and *in vivo* acute toxicity, i.e. when specific organ toxicity is the most sensitive parameter for acute toxicity.

Biokinetic modelling

During the past decades the possibilities to estimate the concentrations of compounds in the different organs and tissues has been greatly improved by the application of simulation models describing these concentrations over time. These so-called physiologically-based biokinetic (PBBK) models (also described as pharmacokinetic models: PBPK) have proven to be very useful in many areas of toxicology and risk assessment (Andersen, 1991; 2003). In essence, these models enable the estimation of the concentration-time relationship of compounds, making use of two essential blocks of information: 1) the known anatomy and physiology of the organism; 2) the physico-chemical properties of the compound. The models describe the relevant anatomical structures such as liver or kidney, or tissue types such as fat or muscle. Tissue volumes and blood flow rates are important parameters in these descriptions. Distribution of a compound throughout the body is described by tissue–blood PCs and, if necessary, by any active transport or biotransformation processes. These species-specific and compound-specific parameters form part of a set of differential equations that describe the biokinetic behaviour of a compound. The feasibility of this modelling approach has been greatly increased by the availability of computer techniques that allow for the simultaneous, numerical solution of differential equations (Clewell and Andersen, 1986).

Using these models, it is possible to extrapolate from one exposure route to another by re-parameterising only the relevant uptake process, as the description of the distribution and elimination processes remains valid. Beside route-to-route extrapolation, the biologically based approach of PBBK models also allows extrapolation of dose and animal species beyond the conditions of laboratory studies (Andersen, 1991).

Biokinetic modelling based on *in vitro* and other non-animal data

The quality of biokinetic models heavily relies on the accuracy of the parameters used to build the models. Many models described in the literature are, at least for the major part, based on parameters measured in *in vivo* experiments. These experiences have resulted in the accumulation of data on parameters describing the physiology of organisms. Literature data now exist for a number of species, including laboratory rodents and humans. These data not only describe the "standard" rat or the "standard" human, but may also allow taking into account the known variability in these parameters, e.g. body weight. On this basis, a number of other parameters can be scaled as a function of body weight, e.g. organ volumes, blood flows, etc. Thus, it is possible to parameterise the models for their physiological parameters in a flexible way without doing experiments.

This is not so obvious for the compound-related parameters. Absorption via the oral, inhalatory or dermal route relies not only on factors determined by the anatomy and physiology of the organism, but also on the physico-chemical characteristics of the compound, and – most importantly – on the interplay between these characteristics and the physiology. The same applies to the distribution of the compound over the organs and tissues as well as to the elimination (excretion, biotransformation).

Attempts have been made to quantify these parameters on the basis of physico-chemical data. This has been done successfully for a number of these parameters, e.g. blood-air and blood-fat partitioning can be described well using quantitative property-property relationships (QPPRs) based on volatility and lipophilicity (Poulin and Krishnan, 1996a; Poulin and Krishnan, 1996b; DeJongh et al.,1997). Other processes are more difficult to describe. QPPRs have been developed for dermal absorption, but these relationships are often limited to certain chemical classes (Wilschut et al., 1995). The same applies for distribution to other tissues, such as the central nervous system, and also for excretion processes and for biotransformation. One important cause is the existence of transport and biotransformation processes that are not easily quantifiable on the basis of structure and physico-chemical properties alone. In these areas further experimental information is needed. The development and validation of *in vitro* methods in these areas is highly desirable (Coecke et al., 2005). In a number of these areas progress is being made, e.g. *in vitro* methods to determine oral absorption, making use of Caco-2 cell systems (Artursson and Borchardt, 1997), co-cultures to quantify blood-brain barrier transport (Prieto et al., 2004), hepatocyte cultures to measure biotransformation parameters (Treijtel et al., 2004).

Extrapolating *in vitro* toxic concentrations to *in vivo* toxic doses: integration of *in vitro* derived toxicodynamic data and biokinetic modelling

A number of studies have attempted to improve the predictive power of *in vitro* cytotoxic concentrations to estimate *in vivo* toxic doses, making use of biokinetic models for which the parameters were derived from *in vitro* and other non-animal data. In the ECITTS programme, the neurotoxicity of 10 chemicals was estimated on the basis of *in vitro* neurotoxic concentrations and biokinetic modelling. For the majority of these compounds, good predictions could be made for their lower-effect levels found in the literature (DeJongh et al., 1999; Blaauboer et al., 2000; Forsby and Blaauboer, 2003).

Another study applying this integration of *in vitro* cytotoxicity and biokinetic modelling was recently published by Gubbels-van Hal et al. (2005). In this study, six endpoints, among them acute toxicity for fish and for rodents, were estimated and compared with traditionally determined (*in vivo*) data from the same laboratory. This parallel testing gave the same acute toxicity classification for all compounds for fish and for 9 out of 10 compounds for rodents.

These studies form the "proof of concept" for the approach of integration of data derived from different domains in toxicolog-

ical hazard and risk assessment, as proposed earlier (Blaauboer et al., 1999; Health Council of the Netherlands, 2001).

The relevant *in vitro* toxic concentrations

One prerequisite for making useful extrapolations from *in vitro* toxic concentrations to *in vivo* toxic doses is that the proper concentrations are taken into account. *In vitro* systems consist of a biological component (i.e. the cell culture), a medium component (with or without serum or other proteins) and a physical component (the test tube or culture plate). Adding a certain amount of compound to the medium will result in a certain concentration to which the cells are exposed. This concentration is highly dependent on the characteristics of the different components in the system, as well as on the processes occurring at the borders of the components (e.g. binding to plastic, evaporation, etc.). Therefore, it is necessary to pay attention to these characteristics and processes. Taking the "nominal" concentration as a basis for further steps in the above-described integrative approach might lead to erroneous extrapolations. Sources of error are: protein binding (Vaes et al., 1997; Seibert et al., 2002; Heringa et al., 2004), the amount of medium in relation to the amount of cells, as well as the aforementioned binding to plastic and evaporation.

These considerations have led to the development of the concept of "*in vitro* biokinetics".

Improving the correlation between *in vitro* cytotoxicity data and acute toxic doses: alerts and correctors for biokinetic behaviour

The main objective of the ACuteTox programme is to produce a reliable system in which the acute toxicity of a compound can be estimated on the basis of non-animal data. The starting point is the determination of basal cytotoxicity. The finding that this parameter is not always a good predictor for acute toxicity *in vivo* stresses the need for the development of a logical and transparent strategy. In such a strategy decisions are made on the requirement of more data than basal cytotoxicity alone. In other words: when do we need to know more about a compound's biokinetic behaviour or its specific organ toxicity and how can we use this knowledge for the estimation of *in vivo* acute toxicity?

When confronted with a newly introduced compound of which the chemical and spatial structure is known, what are the means by which we can be alerted, without further experimentation than standard *in vitro* cytotoxicity testing, to it displaying potential for *in vivo* acute oral toxicity different to what might be expected from its *in vitro* cytotoxicity?

Questions to be answered are then:

- Can we expect that the compound will be absorbed efficiently via the relevant route of exposure? If a compound has a very low rate of absorption, this will lead to low internal exposure.
- Is it to be expected that the compound will have a distribution pattern *in vivo* that will lead to higher or lower concentrations near target cells? The answer to this question is important for the interpretation of basal cytotoxicity data, but can also be used to find clues for organ-specific toxicity.
- Is it to be expected that a compound would be metabolised to a compound with either a higher or with a lower toxicity?
- Can we expect that a compound will have a specific toxic effect that would cause acute toxicity? In that case, basal cytotoxicity would probably underestimate the *in vivo* acute toxicity.
- For which compounds will it be necessary to take *in vitro* biokinetics into account? This question refers to the relevance of a "nominal" concentration in the *in vitro* system vs. the "free" concentration.

From our earlier experiences, as described in the previous paragraphs, it is clear that tools to answer a number of these questions have been developed or are in the process of development. This enables further honing of a strategy. A first step in such a strategy would be to find those parameters in the structure of the compound or in its physico-chemical characteristics that can act as alerts for further action. A further step would include the use of knowledge on comparable or similar compounds (read across). Several papers on the relation between molecular properties and ADME-properties of compounds (QPPRs) (Poulin and Krishnan, 1996a; Poulin and Krishnan, 1996b; DeJongh et al., 1997) and even commercial software packages exist that give answers to at least one aspect of each of these questions.

Based on the estimated ADME-properties, a PBBK model can be developed and free plasma concentrations can be calculated. If one assumes that the *in vitro* culture medium free concentration represents this concentration, corresponding LD_{50} estimations follow. When a metabolite is the toxic agent, its DME properties should be estimated from its physicochemical properties and the corresponding PBBK model is developed.

In the realm of biokinetic studies, the emphasis in the ACuteTox programme will be put on the following aspects, since these are considered to be the most sensitive parameters:

- prediction of biokinetic behaviour on the basis of physico-chemical properties
- *in vitro* measurement of parameters relevant for oral absorption and development of QPPRs for these parameters
- *in vitro* measurement of parameters relevant for blood-brain barrier passage and development of QPPRs for these parameters
- *in vitro* measurement of the free concentration of compounds and development of predictors for deviating *in vitro* biokinetic behaviour
- PBBK modelling for those compounds that will be outliers in the simple relationship between *in vitro* LC_{50} and LD_{50} values.

These elements shall form the framework of an important part of a logical and transparent strategy for the estimation of acute toxicity values.

References

Andersen, M. E. (1991). Physiological modelling of organic compounds. *Ann. Occup. Hyg. 35*, 309–321.

Andersen, M. E. (2003). Toxicokinetic modeling and its applications in chemical risk assessment. *Toxicol. Letters 138*, 9-27

Artursson, P. and Borchardt, R. T. (1997). Intestinal drug absorption and metabolism in cell cultures: Caco-2 and beyond. *Pharmacological Research 14*, 1655-1658.

Blaauboer, B. J., Barratt, M. D. and Houston, J. B. (1999). The integrated use of alternative methods in toxicological risk evaluation. ECVAM Integrated Testing Strategies Task Force Report 1. *Alternatives to Laboratory Animals 27*, 229-237.

Blaauboer, B. J. (2002). The necessity of biokinetic information in the interpretation of in vitro toxicity data. *Altern. Lab. Anim. 30, suppl. 2*, 85-91.

Blaauboer, B. J. (2003). The integration of data on physicochemical properties, in vitro-derived toxicity data and physiologically based kinetic and dynamic modelling as a tool in hazard and risk assessment. A commentary. *Toxicol. Letters, 138*, 161-171.

Blaauboer, B. J., Forsby, A., Houston, J. B. et al. (2000). An integrated approach to the prediction of systemic toxicity using biokinetic models and biological in vitro test methods. In M. Balls, A. M. van Zeller and M. E. Halder (eds.), *Progress in the Reduction Refinement and Replacement of Animal Experimentation* (525–536). Amsterdam: Elsevier.

Clemedson, C. and Ekwall, B. (1999). Overview of the Final MEIC Results: I. The In Vitro-In Vitro Evaluation. *Toxicol. in Vitro 13*, 657-663.

Clewell, H. J. and Andersen, M. E. (1986). A multiple dose-route physiological model for volatile chemicals using ACSL/PC. In F. D. Cellier (ed.), *Languages for Continuous Simulation* (95). San Diego: Society for Computer Simulation.

Coecke, S., Blaauboer, B. J., Elaut, G. et al. (2005). Toxicokinetics and metabolism. *Altern. Lab. Anim. 33 Suppl. 1*, 147-175.

DeJongh, J., Nordin-Andersson, M., Ploeger, B. and Forsby, A. (1999). Estimation of systemic toxicity of acrylamide by integration of in vitro toxicity data with kinetic simulations. *Toxicol. Appl. Pharmacol.158*, 261-268.

DeJongh, J., Verhaar, H. J. and Hermens, J. L. (1997). A quantitative property-property relationship (QPPR) approach to estimate in vitro tissue-blood partition coefficients of organic chemicals in rats and humans. *Arch. Toxicol. 72*, 17-25.

Ekwall, B. (1983). Correlation between cytotoxicity data and LD_{50}-values. *Acta Physiologica et Toxicologica 52 Suppl. II*, 80–99.

Ekwall, B., Barile, F. A., Castano, A. et al. (1998). MEIC evaluation of acute systemic toxicity. Part V. I. The prediction of human toxicity by rodent LD_{50} values and results from 61 in vitro methods. *Altern. Lab. Anim. 26 Suppl. 2*, 617–658.

Forsby, A. and Blaauboer, B. (2003). Integrated in vitro approaches for assessing systemic toxicity. In H. Salem and S. A. Katz (eds), *Alternative Toxicological methods* (43-50). Boca Raton: CRC Press.

Gubbels-van Hal, W. M. L. G., Blaauboer, B. J., Barentsen, H. M. et al. (2005). An alternative approach for the safety evaluation of new and existing chemicals, an exercise in integrated testing. *Regul. Toxicol. Pharmacol., 42*, 284-295.

Halle, W. (2003). The registry of cytoxicity: toxicity testing in cell cultures to predict acute toxicity (LD_{50}) and to reduce testing in animals. *Altern. Lab. Anim. 31 (2)*, 89-198.

Health Council of the Netherlands (2001). Toxicity testing: a more efficient approach. Health Council of the Netherlands. The Hague: Health Council of the Netherlands, 2001; publication no. 2001/24E.

Heringa, M. B., Schreurs, R. H. M. M., Busser, F. et al. (2004). Toward more useful in vitro toxicity data with measured free concentrations. *Environ. Sci. Technol. 38*, 6263-6270.

Poulin, P. and Krishnan, K. (1996a). A tissue composition based algorithm for predicting tissue:air partition coefficients of organic chemicals. *Toxicol. Appl. Pharmacol. 136*, 126-130.

Poulin, P. and Krishnan, K. (1996b). A mechanistic algorithm for predicting blood:air partition coefficients of organic chemicals with the consideration of reversible binding in hemoglobin. *Toxicol. Appl. Pharmacol. 136*, 131-137.

Prieto, P., Blaauboer, B. J., de Boer, A. G. et al. (2004). Blood-brain barrier in vitro models and their application in toxicology. The report and recommendations of ECVAM Workshop 49. *Alternatives to Laboratory Animals 32*, 37-50.

Seibert, H., Mörchel, S. and Gülden, M. (2002). Factors influencing nominal effective concentrations of chemical compounds in vitro: medium protein concentration. *Toxicol. In Vitro 16*, 289–297.

Spielmann, H., Genshow, E., Liebsch, M. and Halle, W. (1999). Determination of the starting dose for acute oral toxicity (LD50) testing in the up and down procedure (UDP) from cytotoxicity data. *Altern. Lab. Anim. 27*, 957-966.

Treijtel, N., Barendregt, A., Freidig, A. P. et al. (2004). Modeling the in vitro intrinsclearance of the slowly metabolized compound tolbutamide determined in sandwich-cultured rat hepatocytes. *Drug. Metabol. Dispos. 32*, 884-891.

Vaes, W. H. J., Urrestarazu Ramos, E., Hamwijk, C. et al. (1997). Solid phase microextraction as a tool to determine membrane/water partition cefficients and bioavailable concentrations in in vitro systems. *Chem. Res. Toxicol. 10*, 1067-1072.

Wilschut, A., ten Berge, W.F., Robinson, P. J. and McKone, T. H. E. (1995). Estimating skin permeation. The validation of five mathematical skin permeation models. *Chemosphere 30*, 1275-1296.

Correspondence to
Bas Blaauboer
Institute for Risk Assessment Sciences
Utrecht University
P.O. Box 80.176
3508 TD Utrecht
The Netherlands
e-mail: b.blaauboer@iras.uu.nl

ACuteTox – Optimation and Pre-validation of an *In Vitro* Test Strategy for Predicting Human Acute Toxicity

Cecilia Clemedson[1], Bas Blaauboer[2], José Castell[3], Pilar Prieto[4], Leila Risteli[5], Joan-Albert Vericat[6] and Albrecht Wendel[7]

[1]Expertrådet Environmental Competence AB, Sollentuna, Sweden; [2]Institute for Risk Assessment Sciences, Utrecht University, The Netherlands; [3]University Hospital La Fe, Valencia, Spain; [4]ECVAM, Institute for Health and Consumer Protection, Joint Research Center, European Commission, Ispra (Va), Italy; [5]Research and Innovation Services, University of Oulu, Finland; [6]NeuroPharma S.A., Preclinical Development, Tres Cantos, Spain; [7]Biochemical Pharmacology, University of Konstanz, Germany

Summary

ACuteTox is an integrated project under the EU-FP6 with the aim to develop a simple and robust in vitro testing strategy for prediction of human acute systemic toxicity, which could replace animal tests used for regulatory purposes.

Studies show good correlation of over 70% between in vitro basal cytotoxicity and rodent LD_{50} values or human lethal blood concentrations. However, a number of discrepancies occur which result in misclassification. ACuteTox aims to identify factors that can eliminate these misclassifications. The outliers in the in vitro/in vivo correlation will be evaluated in order to introduce further parameters (ADE, metabolism and organ specificity), which might improve the correlation. Integration of alerts and correctors in a prediction algorithm, together with implementation of medium throughput approaches, would allow establishment of a new testing strategy to better predict toxic classification.

Keywords: acute toxicity, kinetics, in vitro, in silico, testing strategy, target organ toxicity

Introduction

Validated alternative test methods are urgently required for toxicological safety testing of drugs, chemicals and cosmetics. Both REACH (Registration, Evaluation, Authorisation and Restriction of Chemicals) and the 7th amendment of the Cosmetics Directive (deadline for acute toxicity is 2009) call for the broad replacement of animal experiments on a short-term basis. Animal tests for topical toxicity have been successfully replaced one by one by alternative methods. In systemic toxicity, on the other hand, only some reduction or refinement methods have been developed. The aim of ACuteTox, which started on the 1st of January 2005, is to develop a simple and robust *in vitro* testing strategy to predict human acute systemic toxicity, which could replace the animal acute toxicity tests used today for regulatory purposes.

The extensive amount of work performed since the 70's produced a great number of *in vitro* models for acute systemic toxicity. Many studies have shown a relatively good correlation between *in vitro* basal cytotoxicity data and rodent LD_{50} values. In addition, the MEIC (Multicenter Evaluation of *In vitro* Cytotoxicity) programme showed a good correlation (around 70%) between *in vitro* basal cytotoxicity data and human lethal blood concentrations. However, this correlation means that a certain number of misclassifications have to be faced when using the existing tests. ACuteTox aims to improve this correlation to a level sufficient to ensure a valid prediction of acute toxicity.

ACuteTox is based on two major ongoing activities in this field, i.e. the ECVAM/ICCVAM validation study for basal cytotoxicity tests and the EDIT (Evaluation-guided Development of *In vitro* Test batteries) programme (the continuation of the MEIC study).

The MEIC and the EDIT programmes

The aim of the MEIC study, directed by Björn Ekwall, was to evaluate the relevance of using *in vitro* toxicity tests to predict human acute systemic toxicity (Clemedson et al., 1996a, 1996b, 1998a, 1998b and 2000; Ekwall et al., 1998a, 1998b and 2000).

In summary, an average EC_{50} of ten 24 h exposure tests with human cell lines predicted human peak concentrations from LC_{50} curves better (R^2=0.74) than the prediction of human lethal doses by LD_{50} for rats and mice (R^2=0.60-0.66) for 50 reference chemicals. When some known human toxicokinetic data (knowledge of the passage across the blood-brain barrier (BBB) and the timing of the lethal action) was used together with the cytotoxic concentration to predict human lethal concentrations, prediction increased considerably. The results demonstrated a high relevance of using human cell tests to predict human acute toxicity of chemicals, but showed also that other important toxic mechanisms exist, which may only be measured by supplementary *in vitro* toxicity tests, and that modelling of human toxicity was improved by additional toxicokinetic data, which can probably be obtained by new *in vitro* kinetic tests. These results encouraged Björn Ekwall to initiate the EDIT-project in 1998 (Ekwall et al., 1999). The main aim of the EDIT programme was

to optimise the original MEIC test battery by establishing and validating new *in vitro* tests relevant to biokinetics and tests for organ-specific toxicity.

The ECVAM-ICCVAM validation study of two *in vitro* basal cytotoxicity tests

During the International Workshop on *In vitro* Methods for Assessing Acute Systemic Toxicity (ICCVAM, 2001a), recommendations were made on the need to evaluate two sufficiently developed and standardised basal cytotoxicity assays for their ability to predict rodent LD_{50} values by using the RC (Register of Cytotoxicity) regression model in order to improve the dose selection for *in vivo* studies, and to predict human lethal concentration (ICCVAM, 2001b). In 2002 ECVAM and ICCVAM designed and started a joint validation study. 72 chemicals, which comprise the majority of the 50 MEIC chemicals, and chemicals nominated by different regulatory authorities, were selected according to the availability of acute oral rodent and human toxicity data. These chemicals have been tested in the BALB/c 3T3 cell line and in normal human keratinocytes using the neutral red uptake assay. A preliminary analysis of 12 chemicals showed a good correlation for both cell types between *in vitro* IC_{50} values and human peak concentrations from LC_{50} curves (Casati et al., 2005).

Methods

The starting point for ACuteTox is the existing correlation of *in vitro/in vivo* data (about 70%). Efforts will be dedicated to improve the correlation. This will be done by evaluating outliers of the *in vitro/in vivo* correlation in order to introduce further parameters, such as ADE (administration, distribution and elimination), metabolism and organ specificity, which might improve the correlation. Establishment of a new testing strategy with a better prediction of toxic classification will be possible by integration of alerts and correctors in a prediction algorithm, and by implementation of medium throughput approaches. "Alerts" refer to test/modelling indicating that the given chemical might deviate from the correlation and thus the prediction cannot be

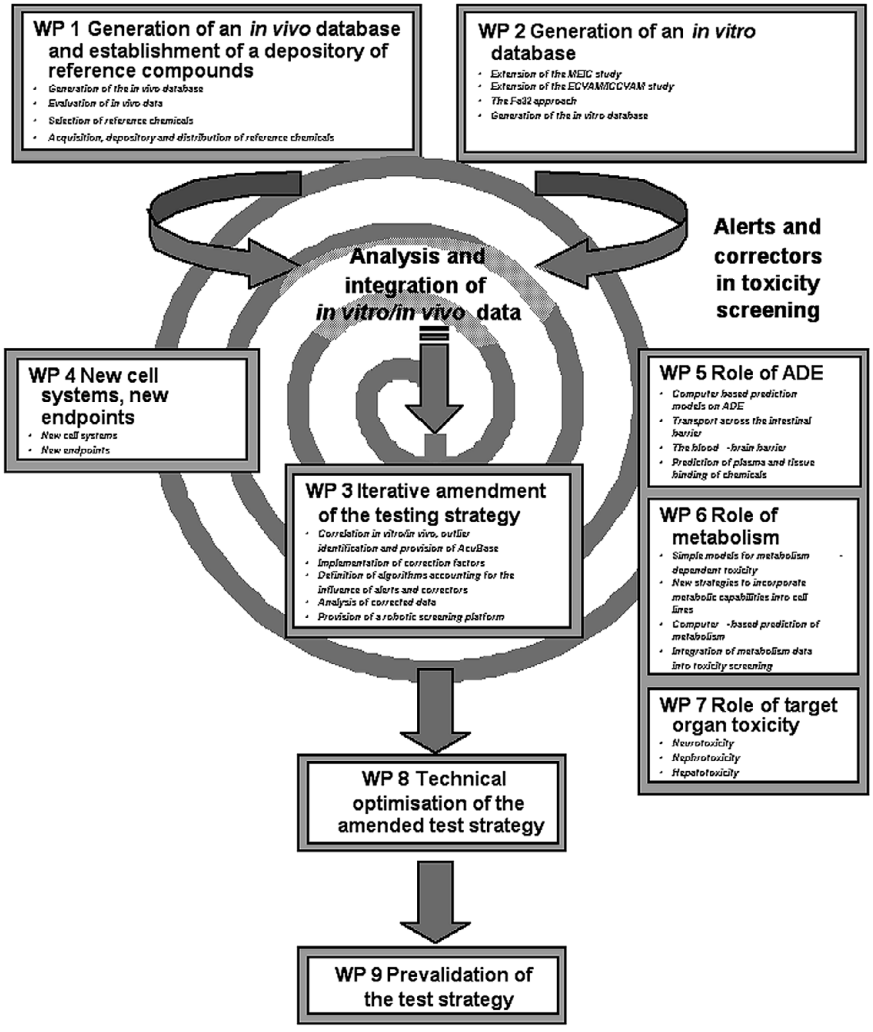


Fig. 1: The structure of the ACuteTox project.

trusted. "Correctors" refer to correction factors amending the result of the prediction.

The work within the project is divided into nine work packages (WPs). Figure 1 shows the outline of the project and the work in the different WPs is summarised below.

Work package 1: Generation of an *in vivo* database

The major objective of this WP is to generate a database containing high quality human and animal data. The *in vivo* databases obtained in the MEIC and ECVAM/ICCVAM studies, containing *in vivo* toxicity data for about 90 chemicals, will be merged to a new database. In total, data for 100-140 chemicals will be gathered. The data will be used in the evaluation of *in vitro* data obtained in WP2-7 and finally in the evaluation of the *in vitro* testing strategy. The animal tests used today in acute toxicity for regulatory purposes have not been properly validated. Therefore, some biometrical retrospective studies with the aim to evaluate the quality and predictive capabilities of the *in vivo* data will be performed. The outcome of these studies will give the reference of how accurate the ultimate *in vitro* test strategy has to be in order to be acceptable for regulatory purposes. Another aim of this WP is to select relevant reference chemicals for the pre-validation study in WP9.

Work package 2: Generation of an *in vitro* database

The aim of this WP is to merge cytotoxicity data from two ongoing projects (ECVAM/ICCVAM and MEIC/EDIT) and to increase the new database by extending the two studies. New testing will be performed in order to get a complete set of cytotoxicity data for the reference chemicals of the MEIC and ECVAM/ICCVAM studies. In addition, a reduced number of compounds showing specific organ toxicity (neuro-, nephro- and hepatotoxicants) have been selected and will be tested. The data will be generated in human primary keratinocytes, human (HepG2, HL-60), mouse (Balbc/3T3) and rat (Fa32) cell lines with different endpoints (protein content, ATP content, and neutral red uptake). The final outcome of WP2 will be a database containing high quality *in vitro* toxicity data obtained with defined Standard Operating Procedures.

Work package 3: Iterative amendment of the testing strategy

The aim of this WP is to reduce step by step the outliers in the *in vitro/in vivo* correlations (data from WP1 and WP2) by examining the underlying reasons for deviation such as the quality of data and lack of ADME or target organ specificities. For this purpose, the WP will evaluate new cell systems and/or endpoints (WP4), and introduce, in an iterative manner, corrector and alert assays for ADE, metabolism and target organ toxicity (WP5-7). This requires detailed characterisation of the outliers that are generated during the process. This iterative approach will allow, in a sequential manner, improvements of the prediction for acute toxicity. During the iterative process, the assays that are identified as THE candidates for corrector and alert assays, based on the improved ability to model toxic and lethal doses, will be adapted and transferred to a robotic system.

Work package 4: New cell systems and new endpoints

The aim of this WP is to provide an alternative way to improve the predictivity of cell-based cytotoxicity assays by incorporating more specific end-point parameters, and/or more appropriate cell systems. Cord blood cells and subpopulations thereof enriched in stem cells and human cell lines will be used as novel *in vitro* models for haematopoietic toxicity assays; colony forming unit-granulocyte/macrophage (CFU-GM), CFU-megakarocyte, and whole blood cytokine production assays. The feasibility of cytomic analysis adapted to robotic screening will be explored by the use of the newest developments and strategies to investigate cytotoxicity changes, by flow cytometric and biochemical assays. Also, endpoints such as oxidative stress and delayed toxicity will be evaluated in different cell systems.

Work package 5: Alerts and correctors in toxicity screening (I): Role of ADE

In this WP, the most crucial parts of the kinetic behaviour will be studied. This will be done, either by experimentally determining kinetic parameters by *in vitro* methods, or by computer-based kinetic modelling. The most crucial parts are: absorption of compounds, distribution between blood and tissues and the passage of special barriers. In the context of acute toxicity, the BBB is the most relevant special barrier and will receive extra attention. The value of *in vitro* determinations of toxic effects will further increase when kinetic parameters are taken into consideration. This will be done by modelling the *in vivo* kinetic behaviour of a compound, making use of physiologically-based biokinetic models (PB-BK).

Work package 6: Alerts and correctors in toxicity screening (II): Role of metabolism

To determine whether toxicity is associated with the metabolism of a compound, cytotoxicity in primary rat hepatocytes (metabolically competent cell) versus HepG2 (metabolically non-competent cell line) will be examined. By comparing the concentration-toxicity curves (or IC_{50} data) of the compound in both models it should be possible to ascertain whether the molecule elicits toxicity after or irrespective of its metabolism, indicating if bioactivation of the xenobiotic is required to elicit toxicity. A subset of reference chemicals, containing compounds for which metabolism is known to be involved in the toxicity mechanism and compounds for which toxicity is independent of metabolism, will be tested. Strategies based on engineered cells, including expression vectors for transient and controllable expression of biotransformation enzymes for CYP 1A2, 2A6, 2C9, 2E1 and 3A4, could be a way to overcome their intrinsic limitations by generation of metabolically competent cell lines. For this purpose, recombinant-defective adenoviral vectors encoding major CYP genes involved in foreign compound metabolism will be generated and used in HepG2 cells.

Work package 7: Alerts and correctors in toxicity screening (III): Role of target organ toxicity

The aim of WP7 is to explore the role of specific organ toxicity (liver, kidney and nervous system) in responses induced by out-

liers in the *in vitro/in vivo* correlation. Since the scientific expertise for each of the organs is very different, this WP is organised into three sub-WPs.

WP 7.1 Neurotoxicity

Lack of correlation between *in vivo* and *in vitro* data can arise from specific effects on main organs, which results in death of the living being without general cytotoxic effects. The vital functions of the central (CNS) and the peripheral nervous system (PNS) suggest these organs as candidates to explain the presence of outliers. Studies of the specific and general targets *in vitro* require well-characterised and complementary model systems. This WP therefore engages organotypic slice cultures, aggregating brain cell cultures, primary cell cultures and neuronal cell lines (also models for the PNS) to assess the effectiveness of *in vitro* models for acute neurotoxicity of a set of selected reference chemicals, including both specifically neurotoxic and general acutely toxic substances.

The major molecular mechanisms involved in stimulatory and inhibitory/depressive activity of the CNS and PNS will be explored. The following endpoints will be studied: $GABA_A$ receptor function, voltage-gated calcium channel function, GABA uptake, acetylcholine esterase-, cholineacetyl transferase-, glutamate decarboxylase-, glutamine synthase- and 2',3'-cyclic nucleotide phosphohydrolase activity, glial and neuronal differentiation, dendrite and axonal structures, Ca^{2+} homeostasis, total ATP content, mitochondrial membrane potential (MMP) and cell membrane potential (CMP), and production of ROS (reactive oxygen species). Effects on the global electric activity and glycolytic activity will be determined in complex brain cell aggregates. Investigations of new neurotoxicological target genes by gene array analysis will also be enrolled. The general cytotoxicity, measured as cell membrane disruption and cell death will be determined in each cell/organotypic model as a reference for the neurospecific effects.

WP 7.2 Nephrotoxicity

It is important to know whether the outliers in the existing correlation between *in vitro* and *in vivo* data present specific nephrotoxic effects. The kidney is especially susceptible to toxicity because of its role in excreting compounds, which involves a high blood supply, concentrating substances, metabolising substances and transporting substances. A major focus of this sub-WP will be the further development of *in vitro* models by analysis of a subset of nephro- and non-nephrotoxic compounds by means of trans-epithelial, paracellular permeability in the LLC-PK1 and MDCK renal tubular cell lines.

WP 7.3 Hepatotoxicity

This sub-WP is aimed at identifying new biomarkers for hepatotoxicity amenable to high throughput testing and capable of alerting to compounds that might show preferential or selective toxicity to the liver that would be underestimated by basal cytotoxicity tests. Efforts will be addressed to (1) design a screening strategy for hepatotoxicity amenable to robotic screening; (2) detection of new markers of hepatotoxicity and (3) design *in vitro* models to assess impairment of bile acid and bilirubin transport.

Work package 8: Technical optimisation of the amended test strategy

After data have been collected in the other WPs, full use of these data will be made by integrating the collected knowledge. This will be done using an integrated scheme in which the collected data are interpreted, and decisions are made.

The data incorporated into the integrated scheme will include:

- Physico-chemical properties of compounds (e.g. lipophilicity, reactivity, molecular weight, volatility, etc.), as well as the presence of certain structural alerts for specific types of biological activity. Where available and where possible, this part of the work will also include the application of quantitative structure activity relationships (QSARs).
- *In vitro* toxicity data, primarily basal cytotoxicity. Furthermore, data on selective cytotoxic reactions (i.e. specific disturbances of physiologically relevant processes in differentiated cell types) will be used to make estimates of concentrations leading to toxicity, which are relevant to acute toxicity *in vivo*.
- Kinetic data: estimates of data for absorption, distribution and elimination as well as metabolism. These data are crucial for the interpretation of all *in vitro* toxicity data and will have to be used to relate the relevant concentrations at which toxicity occurs in *in vitro* systems with a dose that will lead to such toxic levels in an intact organism.

The result of this WP will be an estimate of the toxic dose (or the toxic blood concentration) for the chemical under study. These data will be compared with known *in vivo* toxic dose levels from WP1. The integration of all available data will result in better predictions of toxicity. On the basis of the findings, a further developed testing strategy will be described, depicting the minimal requirements for testing with the highest possible accuracy.

Work package 9: Prevalidation of the test strategy

During the last 2 years of the 5-year project the testing strategy will be pre-validated. A detailed assessment of the reproducibility and relevance of each building block composing the strategy will be performed. The protocols will be amended to be sufficiently standardised and optimised to enter a formal validation process. Pre-validation of the most promising models will be carried out according to ECVAM's criteria by 3-4 laboratories. The results will be independently assessed and, if possible, recommendations for the regulatory acceptance of the developed testing strategy will be presented.

Discussion

In summary, ACuteTox aims to improve the prediction of acute toxicity using *in vitro* methods and, at the same time, to signal which compounds require further testing because their acute toxicity cannot be properly predicted. A pre-validated testing strategy and a pre-validated associated prediction model for acute systemic toxicity will be provided. It is expected that further validation will lead to regulatory acceptance and its incorporation into the set of standardised test guidelines for hazard

assessment of chemicals. The proposed testing strategy has the potential to replace EU methods B.1bis and B.1tris in Annex V of Dir 67/548 EEC and consequently the corresponding OECD Test Guidelines 420 and 423, as well as TG 425. This, in addition, will decrease the need for animal testing and will increase the harmonisation of testing protocols, both on EU level and on the global OECD level.

The implementation of REACH will result in the need for a further assessment of up to 30,000 existing chemicals, which are currently marketed in volumes greater than 1 ton p.a. It is estimated that the testing of these existing chemicals will result in the use of around 4 million animals. The costs for safety testing of chemicals are high and will increase tremendously when REACH is implemented in the EU. The optimisation of the *in vitro* testing strategy within the ACuteTox project will contribute to the establishment of less expensive and more scientifically-based safety testing. In addition, it will guarantee the predictivity of the developed testing strategy and will also provide a testing strategy that can be adapted to a robotic system, which, in the end, can deal with the requirements of the new chemicals policy in terms of testing high numbers of chemicals.

References

Casati, S., Hoffmann, S., Strickland, J. et al., (2005), Analysis of the correlation between in vitro cytotoxicity data and acute toxic effects in humans. *Abstract at the 5th World Congress on Alternatives and Animal Use in the life Sciences*, Berlin, Germany, August 21-25, 2005.

Clemedson, C., McFarlane-Abdulla, E., Andersson, M. et al. (1996a). MEIC evaluation of acute systemic toxicity. Part I. Methodology of 68 in vitro toxicity assays used to test the first 30 reference chemicals. *ATLA 24*, 251-272.

Clemedson, C., McFarlane-Abdulla, E., Andersson, M. et al. (1996b). MEIC evaluation of acute systemic toxicity. Part II. In vitro results from 68 toxicity assays used to test the first 30 reference chemicals and a comparative cytotoxicity analysis. *ATLA 24,* 273-311.

Clemedson, C., Barile, A. F., Ekwall, B. et al. (1998a). MEIC evaluation of acute systemic toxicity: Part III. In vitro results from 16 additional methods used to test the first 30 reference chemicals and a comparative cytotoxicity analysis. *ATLA 26*, 91-129.

Clemedson, C., Andersson, M., Aoki, Y. et al. (1998b). MEIC evaluation of acute systemic toxicity: Part IV. In vitro results from 67 toxicity assays used to test reference chemicals 31-50 and a comparative cytotoxicity analysis. *ATLA 26*, 131-183.

Clemedson, C., Barile, F. A., Chesné, C. et al. (2000). MEIC evaluation of acute systemic toxicity. Part VII. Prediction of human toxicity by results from testing of the first 30 reference chemicals with 27 further in vitro assays. *ATLA 28*, 161-200.

Clemedson, C., Nordin-Andersson, M., Bjerregaard, H. F. et al. (2002). Development of an In vitro Test Battery for the estimation of Acute Systemic Toxicity: An outline of the EDIT project. *ATLA 30,* 313-321.

Ekwall, B., Clemedson, C., Crafoord, B. et al. (1998a). MEIC Evaluation of Acute Systemic Toxicity. Part V. Rodent and human toxicity data for the 50 reference chemicals. *ATLA 26*, 571-616.

Ekwall, B., Barile, F. A., Castano, A. et al. (1998b). MEIC evaluation of acute systemic toxicity: Part VI. The prediction of human toxicity by rodent LD50 values and results from 61 in vitro methods. *ATLA 26*, 617-658.

Ekwall, B., Clemedson, C., Ekwall, B. et al. (1999). EDIT-A New International Multicenter Programme to Develop and Evaluate Batteries of In vitro Tests for Acute and Chronic Systemic Toxicity. *ATLA, 27*, 339-349.

Ekwall, B. and Sjörstrom, M. (2000). MEIC evaluation of acute systemic toxicity. Part VIII. Multivariate partial least squares evaluation, including the selection of a battery of cell line tests with a good prediction of human acute lethal peak blood concentrations for 50 chemicals. *ATLA 28*, 201-234.

ICCVAM (2001a). Report of the International Workshop on In vitro Methods for Assessing Acute Systemic Toxicity. ICCVAM Workshop, Arlington, VA, U.S., October 17-10, 2000. National Institute of Environmental Health Sciences, National Institutes of Health U.S. Public Health Service, Department of Health and Human Services, NIH Publication No: 01-4499, The report is available on the Internet: http://iccvam.niehs.nih.gov/.

ICCVAM (2001b). Guidance Document on Using in vitro Data to Estimate in vivo Starting Dose for Acute Toxicity. National Institute of Environmental Health Sciences, National Institutes of Health U.S. Public Health Service, Department of Health and Human Services, NIH Publication No: 01-4500, The report is available on the Internet: http://iccvam.niehs.nih.gov/.

Acknowledgements

The ACuteTox project is supported by an EU-FP6 grant (FP6-LIFESCIHEALTH-2004-512051).

Correspondence to

Cecilia Clemedson, Ph.D.
Expertrådet AB
Rådan 3
191 40 Sollentuna
Sweden
e-mail: cecilia.clemedson@expertradet.se

Cell Culture Models of the Air-Blood Barrier for the Evaluation of Aerosol Medicines

Claus-Michael Lehr, Michael Bur and Ulrich F. Schaefer
Saarland University, Saarbrücken, Germany

Summary:
Cell culture models of pulmonary epithelia offer excellent opportunities to study transport processes of drugs and other xenobiotics across the air-blood barrier, as well as to assess the inhalation safety of new polymers and other chemicals. After adequate characterisation and validation, such systems may be valuable alternatives to inhalation experiments with animals.
We have been evaluating the pulmonary epithelial cell lines Calu-3, 16HBE14o-, CFBE41o- and A549, as well as primary cultures of human alveolar epithelial cells (hAEpC). Typically, cells are grown on permeable filter supports, allowing the formation of monolayers with functional tight junctions and pharmaceutically relevant transporter proteins. While the cell lines Calu-3 and 16HBE14o- appear useful to model the bronchial epithelium, the cell line A549 develops only weak barrier properties. Therefore, it still appears necessary to use primary cultured cells to model the alveolar epithelium.

Keywords: epithelial permeability, cytotoxicity, pulmonary drug delivery, in vitro models

Introduction

The pulmonary route is of increasing interest for the development of new medicines, not only for the treatment of lung diseases (e.g. asthma, COPD) but also for the fast and efficient delivery of drugs into the systemic blood circulation. Advanced drug carriers, such as nanoparticles or liposomes, however, require the use of polymers and other excipients, the effects of which on the airway and respiratory epithelia are still relatively unknown, especially with regard to their safety.

While the safety and efficacy of new drugs and delivery technologies can only be judged by clinical studies in man, such tests are usually preceded by preclinical tests on animals. However, animals by their nature are rather complex systems and do not easily allow the acquisition of detailed information on the mechanisms of drug absorption at a given biological barrier, such as the air-blood barrier of the lung. Therefore, *in vitro* models of biological barriers are extremely useful, because they allow the study of biological processes at such barriers under controlled conditions. Apart from the possibility to use such systems for absorption/safety screening among larger numbers of candidates, studies at the cellular level provide a better understanding of critical factors and therefore allow optimisation of formulations. In the drug development process, where the first aim is to bring a new medicine into clinical testing as quickly as possible, such cell culture based *in vitro* test systems may significantly reduce the use of laboratory animals. In addition, candidate drugs, which have successfully passed the stage of *in vitro* testing, may be expected to have a better success rate in passing the next stages of more complex *in vivo* tests (s. fig.1).

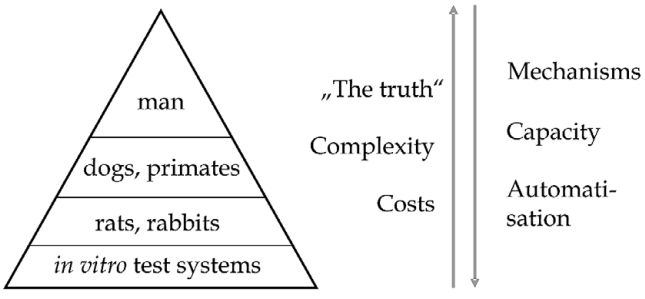


Fig. 1: *In vitro* test systems – the golden mean between manageability and explanatory power.

Functional anatomy of the lung regarding drug absorption and delivery

Some anatomical peculiarities of the human lung have led to an increasing interest in the pulmonary route for both local and systemic drug delivery. The large pulmonary epithelial surface estimated at 100-145 m^2 has almost the same resorption area of the intestinal mucosa, but only a small number of medicines are currently available for inhalation. Also, the high blood perfusion rate and the very small alveolar fluid volume of 7-20 ml are in favour of the lung as a novel route for the application of medicines. The alveolar epithelium in the deep lung is one of the thinnest barriers in the human body. The distance between the airspace and the capillary blood is in the range of 2 and 10 µm and can supposedly also be passed by relatively large molecules. However, not only the alveolar region should be contemplated but also the bronchial region. As conducting airways the bronchial tubes play a fundamental role in drug efficacy, especially for locally acting drugs to treat respiratory diseases such as asthma or COPD. In both the deep lung and

in the bronchial region the epithelium is the key for successful drug delivery.

Transepithelial transport processes
The function of an epithelium is the control of resorption or secretion of substances. Some substances like water can diffuse through the epithelium by using the intercellular space, but most substances need an energy dependent transport. Five main transport pathways are distinct (s. fig. 2): 1) paracellular diffusion; 2) transcellular diffusion; 3) carrier-mediated uptake at the apical domain followed by passive diffusion across the basolateral membrane; 4) transcytosis; 5) active secretion or efflux.

The airway epithelium of the bronchial region
According to the function of the airways – conduction, heating, clearance, and moisturisation of the air – the bronchial region has a specialised epithelium. It consists of columnar shaped cells. In the uppermost areas of the respiratory tract (nose, trachea, bronchi) it is multilayered, but tends to be thinner and monolayered towards the deeper (peripheral) lung. Mucus and ciliary transport are very effective clearance mechanisms. Airway goblet cells and submucosal glands form the major sources of human respiratory mucus. In the adult, mucus-secreting glands occupy about one-third of the inner airway wall wherever there is supportive cartilage.

The respiratory epithelium of the alveolar region
According to its major function, i.e. gas exchange, the alveolar epithelium is very thin (typically less than 1 µm). The largest part (70-90%) of its big surface area (100-140 m^2) is formed by so called alveolar epithelium type I (ATI) cells. They are spread-out in shape and form a tight squamous monolayered epithelium. Interspersed between the ATI cells are the alveolar epithelium type II (ATII) cells. They are relatively small, cuboidal in shape and more numerous than the ATI cells. They occupy only 10-30% of the surface area and are considered to be the progenitors of the ATI cells, because ATII cells may transform into ATI cells if some of the latter die or are damaged. However, the major function of the ATII cells is the production and secretion of lung surfactant.

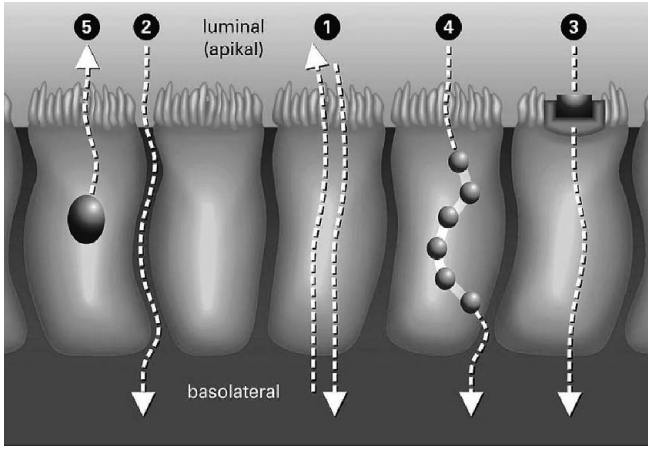


Fig. 2: Transport mechanism of drugs across an epithelium. 1. transcellular diffusion; 2. paracellular diffusion; 3. carrier-mediated uptake at the apical domain followed by passive diffusion across the basolateral membrane; 4. transcytosis; 5. active secretion or efflux. *Deutsche Apotheker Zeitung, 22.01.2004*, 61.

Cell culture models of pulmonary epithelia

Calu-3
Calu-3 is an adenocarcinoma cell line derived from a 25-year old Caucasian male. It has been suggested to express tight barrier properties on the basis of electrophysiological studies. The presence of tight junctional proteins was confirmed by immunoblotting, and functional properties of the monolayers were studied by measurement of transepithelial electrical resistance and mannitol permeability (Wan et al., 2000). Calu-3 cells have been the subject of relatively many investigations. After a few investigations on the equipment of the cells with ion channels or receptors, the cell line was used relatively soon as a tool for transport studies.

Mathia et al. (2002) studied the permeability characteristics of Calu-3 to passively and actively transported drugs and correlated the data with other *in vitro* models and with rat lung absorption *in vivo*. Air-interface cultured (AIC) Calu-3 cells grown on collagen-coated permeable filter supports formed "tight" polarised and well differentiated cell monolayers with apical microvilli and tight-junctional complexes. Solute permeability was dependent on lipophilicity, and inversely related to molecular size. Calu-3 cells actively transported amino acids, nucleosides and dipeptide analogues, but not organic anions, organic cations or efflux pump substrates. The permeability characteristics of Calu-3 cells correlated well with primary cultured rabbit tracheal epithelial cells *in vitro*, and the rate of drug absorption from the rat lung *in vivo*.

Apart from their use as a transport model, Calu-3 cells can also be employed for the investigation of metabolic processes. Borchard et al. (2002) cultivated Calu-3 cells on microporous filters at an air interface for 16-18 days, and incubated the cells with the glucocorticosteroid budesonide. With the aid of mass spectrometry of cell extracts, fatty acid conjugates of budesonide were detected. It seems that Calu-3 cells are able to store budesonide by intracellular conjugation. Therefore, it was suggested to use the Calu-3 cell model as a tool for examination of local pharmacokinetics and metabolism of glucocorticosteroids at the bronchial epithelium. Glucocorticosteroids were also employed for a study on the efflux system P-glycoprotein in Calu-3 cells (Hamilton et al., 2001). The P-gp modulation efficacy of glucocorticosteroids was determined by its ability to increase the accumulation of the P-gp substrate rhodamine 123 in the cells. Because of the high tightness and the easy cultivation conditions, Calu-3 cells are widely used for transport studies. Although Calu-3 is a bronchial (i.e. not alveolar!) epithelial cell line, it is often used also as a model of the pulmonary epithelium in general.

16HBE14o-

Another human bronchial epithelial cell line, 16HBE14o-, immortalised by virus transformation, also shows significant transepithelial resistance and can be used for transport studies. In comparison to the Calu-3 cells, the 16HBE14o- cell line seems to express more P-glycoprotein, lung resistance-related protein (LRP), and caveolin-1. Immunocytochemical staining showed expression of P-gp localised at the apical membrane of 16HBE14o- cell layers. The flux of rhodamine 123 across cell layers exhibited a greater apparent permeability (P_{app}) value for the secretory direction. This asymmetry disappeared in the presence of verapamil, a P-gp inhibitor. The 16HBE14o- cell line may be a suitable candidate for an *in vitro* model for mechanistic studies of drug transport processes involved in the smaller airways, because it shows drug transport systems that are also present in the human bronchus *in vivo* (Ehrhardt et al., 2003). To increase the simulation of the *in vivo* conditions, the cultivation of the cell lines under air-interface culture conditions was attempted. ZO-1, as indicator for the tight junctions, was found in cells grown in both AIC and LCC (liquid culture conditions). However, only LCC-grown cells exhibit protein ZO-1 localised as a zonula-occludens-like regular belt connecting neighbouring cells. The presence of typical tight junctions has been confirmed by electron microscopy. Immunostaining for occludin, claudin-1, connexin 43, and E-cadherin has demonstrated intercellular junction structures only in the cells in LCC. These morphological findings were paralleled by higher transepithelial electrical resistance values and similar fluxes of the hydrophilic permeability marker fluorescein-Na under LCC compared with AIC conditions (Ehrhardt et al., 2002). Because of the equipment of the cell with drug transport systems like P-gp or LRP, this cell line can be used for the investigation of the underlying transport pathways.

Also it has been shown that diesel exhaust particles can be phagocytosed by 16HBE14o- cells, inducing the release of cytokines (Marano et al., 2002).

CFBE41o-

The CFBE41o- cell line was generated by transformation of cystic fibrosis tracheo-bronchial cells with SV40. It is homozygous for ΔF508-CFTR over multiple passages in culture and expresses a number of proteins relevant in the context of pulmonary drug absorption, for example P-gp, LRP and caveolin-1 (Ehrhardt et al., 2005). Cystic fibrosis (CF) is a lethal genetic disease caused by a mutation in the cystic fibrosis transmembrane conductance regulator (CFTR), which mainly functions as a chloride channel. The main clinical symptoms are chronic obstructive lung disease with excessive inflammation and chronic infection, which is responsible for most of the morbidity and mortality associated with CF, and pancreatic insufficiency. The CFBE41o- cell line should be useful for studies in the scope of CF gene transfer or alternative treatment using small drug molecules and gathering further knowledge about the disease on the cellular level, without the need for primary culture.

A549

The A549 cell line possesses alveolar epithelium type II cell phenotype and has been widely used as a system to study the regulation of pulmonary surfactant synthesis. However, cultured A549 cells do not undergo transition to form a phenotype similar to that of an ATI cell. Furthermore, although the A549 cell has received some attention as a monolayer culture for the study of solute transport, its cell architecture and barrier properties are quite distinct from that of an ATI cell monolayer. Thus, an *in vitro* cell model of the human alveolar epithelium possessing the relevant qualities of the alveolar epithelium *in situ* is definitely needed. The A549 line is a human lung adenocarcinoma derived by explant culture from the peripheral airways of a Caucasian male with lung cancer. A549 cells show a very high mannitol permeability coefficient, and approach the characteristics of cell-free filters. The "leaky" monolayers formed by these airway carcinoma cell lines failed to show significant immunostaining for the tight junction protein ZO-1 (s. fig. 3). The leaky formation of tight junctions in A549 is also the cause of very low transepithelial electrical resistance by their monolayers. This suggests that the formation of peripheral rings of ZO-1 staining is related to the formation of tight junctions and that these junctions probably determine low permeability to mannitol. A549 exhibited staining for desmoplakin, but no staining for E-cadherin. The functional tight junction deficits of the A549 cell line seem to preclude its use in permeability studies. Nevertheless, some authors have also reported comparatively high TEER (transepithelial electrical resistance) values and permeability rates for filter-grown A549 cells, which might be a question of optimised culture conditions

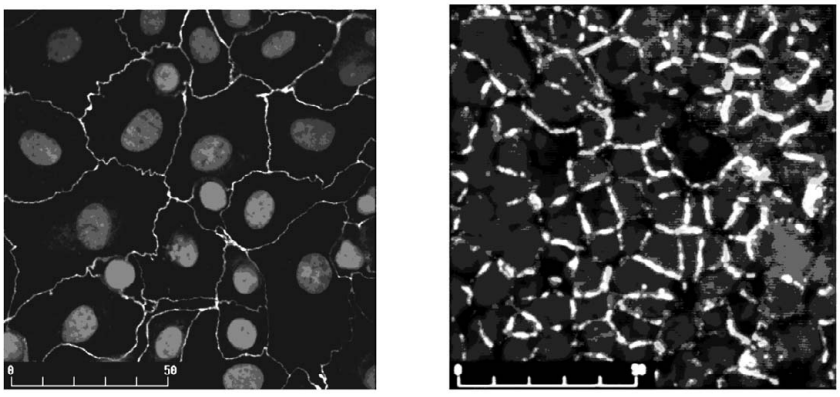


Fig. 3: Visualisation of tight junctional protein ZO-1 by specific antibody; nuclei counterstained by propidium iodide and observed by confocal laser scanning microscopy.

On the left, human alveolar epithelial cells in primary culture on day 8, and on the right A549 cells after 8 days in culture (Elbert et al.,1999).

(Rothen-Rutishauser et al., 2005). The use of immortalised cell lines is limited by the fact that in many cases the cells loose their characteristics during *in vitro* culture and will senesce after a certain number of cell divisions. Immortal cell lines from primary cultures are not a perfect representation of the original cells in primary culture. Because of these problems a great majority of researchers resort to the use of primary non-cancer cell lines.

Human Alveolar Epithelial cells in primary Culture (hAEpC)
Lung alveolar epithelium *in vivo* is composed of two specialised epithelial cell types, the squamous alveolar epithelial type I (ATI) cell and the surfactant-producing cuboidal alveolar epithelial type II (ATII) cell. Current evidence supports the hypothesis that ATII cells serve as the sole progenitors of ATI cells *in vivo* (Uhal, 1997; Fehrenbach, 2001). Accordingly, isolated ATII cells in culture lose their characteristic phenotype and acquire morphological and biochemical markers characteristic of ATI cells over a 5- to 10-day period. Morphological changes during differentiation include the generation of monolayers with high transepithelial electrical resistance (>1,000 Ohm * cm^2) and a loss of microvilli, an increase in the cell surface area and the development of thin cytoplasmic attenuations extending away from a protruding nucleus. The isolation of ATII cells predominantly from rat and rabbit lung tissue, and their culture over time leading to a primary culture of ATI-like cells is now an established technique for different purposes. Although the isolation of primary human alveolar cells has been described before (Alcorn et al., 1997), primary human cells are not commonly used as an *in vitro* model for the air-blood barrier. The isolation of human alveolar type II epithelial cells (hAEpC) and their primary culture, which results in confluent monolayers capable of generating tight junctional complexes and high transepithelial electrical resistance, was described by Elbert et al. (1999). The morphological cell change from an ATII phenotype to an ATI-like cell phenotype over culture time was described by Fuchs et al. (2003). Moreover, the formation of characteristic plasma membrane structures termed caveolae and the synthesis of their major structural protein, caveolin-1, was observed in these cells. The caveolae membrane system is of interest because of its potentially important role in macromolecule transport across the "air-blood-barrier" of the lung (Gumbleton et al., 2000), including both the clearance of endogenous protein from the airspace and the absorption of inhaled therapeutic protein. Primary type II alveolar cells (ATII) are isolated from human non-tumour lung tissue, which is obtained from patients undergoing lung resection. The isolation is performed according to a protocol described by Elbert et al. (1999). The isolated ATII cells are seeded on collagen/ fibronectin-coated polyester filter inserts using small airway growth medium containing penicillin and streptomycin and with the addition of low concentrations of foetal calf serum in order to suppress fibroblast growth. Formation of functional tight junctional complexes and generation of confluent monolayers is routinely determined by measuring TEER using an electronic voltmeter. After reaching confluence, the alveolar monolayers of hAEpC typically reveal TEER values of 1,000-2,000 Ohm * cm^2 on days 6-8 post seeding. The formation of tight junctions can also be routinely monitored by immunofluorescent staining for ZO-1 (Fuchs et al., 2003).

Pulmonary cell culture models in drug delivery research

Simple test systems are needed for drug absorption studies, especially in the early or exploratory phase of drug development. Excised lung tissue, isolated single cells, artificial membranes and *in vivo* models have distinct limitations. Therefore, the use of filter-grown tight cultures of pulmonary epithelial cells appears to be the most attractive alternative. While the use of similar models, such as e.g. the Caco-2 line, has already found much acceptance with respect to candidate screening for oral drug delivery, the development of comparable models with respect to pulmonary drug delivery is still in its beginnings. Their usefulness with respect to a better understanding of the transport mechanisms (which may also have impact on the design of advanced drug delivery systems) at the air-blood barrier has already been well recognised in the literature. However any prediction of *in vivo* pulmonary drug absorption based on such *in vitro* data can at present only be made most cautiously. In view of the complexity of pulmonary drug deposition (besides absorption!) and its poor control in many clinical studies, there is not enough data available yet to judge or postulate a reliable *in vitro/in vivo* correlation for such pulmonary cell culture systems at the present time.

With this disclaimer in mind, epithelial cell culture models of the lung still bear several advantages over other experimental techniques: (a) they are less time-consuming; (b) they enable rapid evaluation of methods for improving drug absorption; (c) they allow use of human rather than animal tissues; (d) they can help to reduce the number of animal studies.

Absorption and transport studies
Plating tight pulmonary cells on membrane supports, such as Transwell® cell culture inserts, allows the study of apical to basolateral (A → B) and basolateral to apical (B → A) drug transport. Membrane filters have been used as cell growth substrates since the 1950s.

For epithelial and other cell types, the use of permeable supports *in vitro* allows cells to be grown and studied in a polarised state under more natural conditions. Cellular functions, such as transport, adsorption, and secretion can also be studied, since cells grown on permeable supports provide convenient, independent access to apical and basolateral plasma membrane domains. Through taking samples from the receiver compartment, drug transport can be measured. The permeability of the investigated compound can be calculated by

$$P_{app} = (dQ/ dt)/ (A* C_0),$$

where dQ/ dt (mol s^{-1}) is the transport rate and indicative of the increase in the concentration of drug in the receiver chamber per

time interval. A (cm^2) is the surface area of the cell culture support, and C_0 (mol/ml) the initial drug concentration in the donor chamber.

Safety assessment/cytotoxicity
Inhalable nanoparticles are able to enhance drug or DNA stability for purposes of optimised deposition to the targeted lung areas. Surface modifications can mediate drug targeting. The suitability and non-toxicity of nanoparticles and the raw material of the nanoparticles have to be investigated *in vitro* on primary airway epithelium cells and cell lines like 16HBE14o- or Calu-3. The uptake of nanoparticles into these cells can be examined by confocal laser scanning microscopy or flow cytometry. The cytotoxicity of particles can be evaluated by a LDH (lactate dehydrogenase) release test and the inflammatory potential of the particles can be assessed by measuring interleukin (IL-8) release. Brzoska et al. (2004) for example reported with the aid of a 16HBE14o- model that protein-based nanoparticles are suitable drug and gene carriers for pulmonary application.

Cell compatible deposition systems of metered dose pharmaceutical aerosols
While toxicological studies on aerosols are typically focusing on the exposure to a xenobiotic at a certain concentration and over a given period of time, pharmaceutical aerosols are usually administered as a metered single dose. In order to estimate the bioavailability of a deposited dose, or to judge the biocompatibility of a once or repeatedly administered aerosol formulation, appropriate deposition chambers are needed which are able to address pharmaceutically relevant questions but at the same time are compatible with the implementation of the cell culture systems to study drug absorption subsequent to deposition of aerosol particles.

An *in vitro* model for aerosol deposition and transport across epithelia in the human airways may be a good predictor of therapeutic efficacy. An Andersen viable cascade impactor was used as a delivery apparatus for the deposition of particles onto monolayers of Calu-3 cells (Cooney et al., 2004). It was shown that these cell layers can withstand placement in the impactor, and that permeability can be tested subsequent to removal from the impactor. A simple, multi-stage cascade impinger is also a suitable physiologically-relevant model of the pulmonary epithelial barrier that would allow for quantitative characterisation of therapeutic aerosols *in vitro* (Fiegel et al., 2003). Calu-3 cells grown under air-interfaced culture (AIC) on a semipermeable membrane were impinged with aerosolised large porous particles. These particles were deposited homogeneously and reproducibly on the cell surface and caused no apparent damage to cell monolayers as evaluated by SEM and light microscopy. The monolayers showed no significant change in barrier properties within the first 90 min following particle application. This *in vitro* model, based on AIC grown Calu-3 cells, could allow a more relevant and quantitative characterisation of therapeutic aerosol particles intended for delivery to the tracheobronchial region of the lung or to the nasal passages.

Outlook

The "air-blood barrier" is not only relevant for the safe and efficient delivery of drugs, but is also subject to exposure to various air-borne pollutants in the form of (nano) particles. Therefore, the lung plays a major role as a target organ in any systemic toxicity test (acute, subchronic, and chronic) as well as in inhalation toxicology or respiratory sensitisation. In all these areas suitable alternatives to animal experiments are desperately sought.

References

Alcorn, J. L., Smith, M. E., Smith, J. F. et al. (1997). Primary cell culture of human type II pneumonocytes: maintenance of a differentiated phenotype and transfection with recombinant adenoviruses. *Am. J. Respir. Cell Mol. Biol. 17*, 672-682.

Borchard, G., Cassara, M. L., Roemele, P. E. et al. (2002). Transport and local metabolism of budesonide and fluticasone propionate in a human bronchial epithelial cell line (Calu-3). *J. Pharm. Sci. 91*, 1561-1567.

Brzoska, M., Langer, K., Coester, C. et al. (2004). Incorporation of biodegradable nanoparticles into human airway epithelium cells-in vitro study of the suitability as a vehicle for drug or gene delivery in pulmonary diseases. *Biochem Biophys. Res. Commun. 318*, 562-570.

Cooney, D., Kazantseva, M. and Hickey, A. J. (2004). Development of a size-dependent aerosol deposition model utilising human airway epithelial cells for evaluating aerosol drug delivery. *Altern. Lab. Anim. 32*, 581-590.

Ehrhardt, C., Collnot, E. M., Baldes, C. et al. (2005). Towards an in vitro model of cystic fibrosis small airway epithelium: Characterisation of the human bronchial epithelial cell line CFBE41o-. *Cell Tissue Res. 25*, 1-11.

Ehrhardt, C., Kneuer, C., Fiegel, J. et al. (2002). Influence of apical fluid volume on the development of functional intercellular junctions in the human epithelial cell line 16HBE14o-: implications for the use of this cell line as an in vitro model for bronchial drug absorption studies. *Cell Tissue Res. 308*, 391-400.

Ehrhardt ,C., Kneuer, C., Laue, M. et al. (2003). 16HBE14o- human bronchial epithelial cell layers express P-glycoprotein, lung resistance-related protein, and caveolin-1. *Pharm. Res. 20*, 545-551.

Elbert, K. J., Schaefer, U. F., Schafers, H. J. et al. (1999). Monolayers of human alveolar epithelial cells in primary culture for pulmonary absorption and transport studies. *Pharm. Res. 16*, 601-608.

Fehrenbach, H. (2001). Alveolar epithelial type II cell: defender of the alveolus revisited. *Respir. Res. 2*, 33-46.

Fiegel, J,, Ehrhardt, C., Schaefer, U. F. et al. (2003). Large porous particle impingement on lung epithelial cell monolayers - toward improved particle characterization in the lung. *Pharm. Res. 20*, 788-796.

Forbes, B., Wilson, C. G. and Gumbleton M. (1999). Temporal dependence of ectopeptidase expression in alveolar epithelial cell culture: implications for study of peptide absorption. *Int. J. Pharm. 15 180*, 225-234.

Fuchs, S., Hollins, A. J., Laue, M. et al. (2003). Differentiation of human alveolar epithelial cells in primary culture: morphological characterization and synthesis of caveolin-1 and surfactant protein-C. *Cell Tissue Res. 311*, 31-45.

Gumbleton, M., Abulrob, A. G. and Campbell, L. (2000). Caveolae: an alternative membrane transport compartment. *Pharm. Res. 17*, 1035-1048.

Hamilton, K. O., Yazdanian, M. A. and Audus, K. L. (2001). Modulation of P-glycoprotein activity in Calu-3 cells using steroids and beta-ligands. *Int. J. Pharm. 228*, 171-179.

Marano, F., Boland, S., Bonvallot, V., Baulig, A. and Baeza-Squiban, A. (2002). Human airway epithelial cells in culture for studying the molecular mechanisms of the inflammatory response triggered by diesel exhaust particles. *Cell Biol. Toxicol.*, 315-20.

Mathia, N. R., Timoszyk, J., Stetsko, P. I., Megill, J. R., Smith, R. L. and Wall, D. A. (2002). Permeability characteristics of calu-3 human bronchial epithelial cells: in vitro-in vivo correlation to predict lung absorption in rats. *J. Drug Target. 10*, 31-40.

Rothen-Rutishauser, B. M., Kiama, S. G. and Gehr, P. (2005). A three-dimensional cellular model of the human respiratory tract to study the interaction with particles. *Am. J. Respir. Cell Mol. Biol. 32*, 281-289.

Wan, H., Winton, H. L., Soeller, C., Stewart, G. A., Thompson, P. J., Gruenert, D. C., Cannell. M. B., Garrod, D. R. and Robinson, C. (2000). Tight junction properties of the immortalized human bronchial epithelial cell lines Calu-3 and 16HBE14o-. *Eur. Respir. J. 15*, 1058-1068.

Correspondence to
Prof. Dr. Claus-Michael Lehr
Biopharmaceutics and Pharmaceutical Technology
Saarland University
66125 Saarbrücken
Germany
e-mail: lehr@mx.uni-saarland.de
http://www.uni-saarland.de/fak8/lehr/

Session 5.07
Progress in quality assurance for *in vitro* alternative studies

A Novel Quantitative *In Vitro* Model of Angiogenesis

Mahtab Bahramsoltani, Sabine Käßmeyer and Johanna Plendl
Faculty of Veterinary Medicine, Institute of Veterinary Anatomy, FU Berlin, Germany

Summary
Quantification of angiogenesis, which is done particularly in animal models, is a prerequisite for the determination of the angiogenic or angiostatic effect of substances.
In the frame of this study, a new method for quantification of angiogenesis in vitro was established based on staging of angiogenesis by breaking down the angiogenic cascade into quantifiable steps. This method allows quantification of all phases of angiogenesis up to the development of lumenised, capillary-like structures in vitro. Validation of the method showed that routine and reproducible performance should be possible for different investigators with a maintainable effort of time and costs.

Keywords: angiogenesis, quantification, in vitro model, replacement

Introduction

The process of angiogenesis includes sprouting of new vessels from pre-existing ones, their remodelling and regression. Angiogenesis is a pre-requisite for growth and differentiation of organs and tissues as it allows for supply with nutrients and oxygen (Folkman, 2003). Angiogenesis also plays a key role in many pathological processes, particularly growth and metastasis of tumours (Folkman, 1995; Caldwell et al., 2005; Banning, 2005). Both, the stimulation of angiogenesis to generate new vessels in ischemia and the inhibition of angiogenesis, so called antiangiogenesis, to arrest growth and metastasis of tumours, are promising therapeutical concepts (Gowda, 2005; Markkanen et al., 2005; Huber et al., 2005; Gasparini et al., 2005).

Angiogenesis occurs in a characteristic multi-step cascade of migration, proliferation, differentiation and three-dimensional organisation of endothelial cells to generate new vessels with an internal lumen. Angiogenesis as well as antiangiogenesis are regulated by soluble factors, which may influence endothelial cells in the different stages of the angiogenic cascade (Augustin, 2003).

Quantification of angiogenesis and antiangiogenesis, i.e. their depiction in measurable factors, is a basic requirement for determination of angiogenic or angiostatic effects of a substance. Most quantitative investigations are undertaken in animal models. The most frequently used models are the cornea model, the chorioallantoismembrane (CAM)-model and several skin preparations (Hasan et al., 2004; Bahramsoltani, 2004). It cannot be obviated that these investigations are associated with pain for the animals. Up to now some two- and three-dimensional *in vitro* models of angiogenesis have been developed (for example: Montesano et al., 1983; Nehls and Drenckhahn, 1995; Hoying and Williams, 1996; Meyer et al., 1997; Peters et al., 2002; Rookmaaker et al., 2005; Wang et al., 2005; Pozzi et al., 2005). However, in these models, the effects of the substances tested were quantified in only a few phases of angiogenesis.

The aim of this study was to establish a method to quantify angiogenesis *in vitro* in order to achieve a replacement and complementary method comprising all stages of the angiogenic cascade. Furthermore, routine accomplishment of quantification should be possible for different investigators with a maintainable effort of time and costs.

Materials and methods

Microvascular endothelial cells used were isolated from *corpora lutea* of slaughtered cattle in different development stages (Plendl et al., 1996; Plendl, 1997; Plendl, 2000). Endothelial cells were identified by endothelial markers including the localisation of von Willebrand factor, vascular endothelial growth factor receptors 1 and 2 (VEGFR-1 and 2) and platelet/endothelial cell adhesion molecule (PECAM) (Plendl et al., 1996; Plendl et al., 2002a,b).

Cultivation of endothelial cells was carried out on gelatine (1.5% in PBS, Difco Laboratories, Detroit, USA) coated 24 well plates (Iwaki, Tokyo, Japan). Cells were seeded in a concentration of 41,000 to 45,000 cells per well.

Basic medium included Dulbecco's Modified Eagle's Medium (DMEM) with 10% foetal bovine serum, 1% L-glutamine and 1% penicillin/streptomycin (10,000 U/ml) (all from Sigma-Aldrich, Taufkirchen, Germany).

Endothelial cells were stimulated for angiogenesis using a selective medium, which additionally contained 10% foetal bovine serum, 1% BME-vitamins (Sigma-Aldrich), 1% heparin-solution (0.25%; Sigma-Aldrich), 2% Endothelial Cell Growth Supplement (2.5 mg/ml; Schubert & Weiss, Munich, Germany) und 20% S 180-conditioned medium (see below) in basic medium.

S 180-conditioned medium was produced by cultivation of murine sarcoma cells (courtesy of Prof. R. Auerbach, University of Wisconsin-Madison, WI, USA) in basic medium. Medium was removed and spun down with 1050 U. Supernatant was sterile filtered with 0.2 µm filters (Schleicher & Schüll GmbH, Dassel, Germany) and deep frozen at -20°C until use.

For quantification of angiogenesis digital pictures of defined areas were taken with a video camera (Inteq 000610; Inteq, Berlin, Germany) using the image editing system Axiovision (Version 3.0; Zeiss, Jena, Germany) twice a week. Visual fields were standardised by always taking pictures at the same magnification (100x) and of the same area (691,200 μm^2).

Results

For quantification of all phases of angiogenesis the angiogenic cascade was broken down into quantifiable steps.

Time dependent changes of endothelial cell morphology in the course of the angiogenic cascade were classified into 8 defined stages:

Stage 1: Confluent monolayer
Cells in cobble-stone pattern

Stage 2: Endothelial sprouting, early phase
Sprouting in < 50% of cells

Stage 3: Endothelial sprouting, late phase
Sprouting in > 50% of cells

Stage 4: Linear side-by-side arrangement of cells, early phase
Linear arrangement in < 50% of cells

Stage 5: Linear side by side arrangement of cells, late phase
Linear arrangement in > 50% of cells

Stage 6: Networking of endothelial cells
Network of linearly arranged cells (fig. 1)

Stage 7: Three-dimensional organisation of cells, early phase
Appearance of lumenised capillary-like structures; verification of an internal lumen by electron microscopy (fig. 2). Capillary-like structures are defined as linear structures of endothelial cells with a diameter of more than 28 µm.

Stage 8: Three-dimensional organisation of cells, late phase
All linearly arranged cells form lumenised capillary-like structures

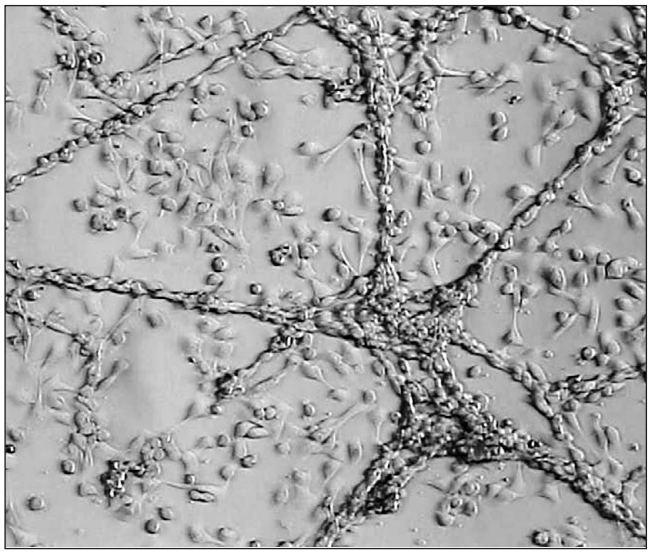


Fig. 1: Endothelial cells from bovine corpus luteum after 40 days in culture. Network of linearly arranged cells. Phase contrast microscopy, 100x.

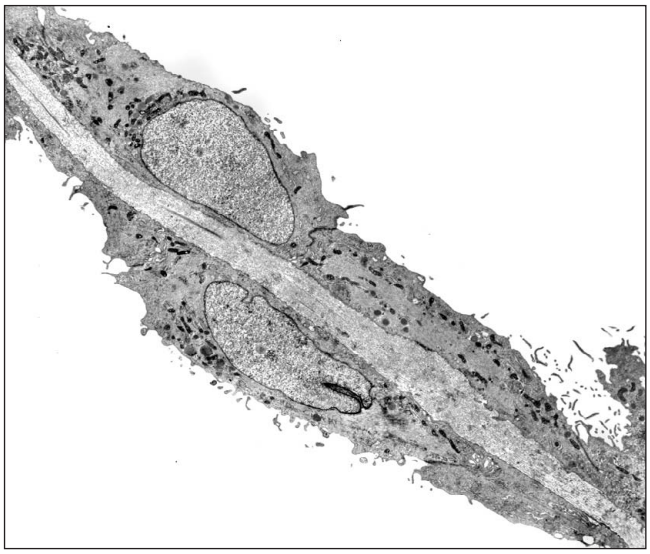


Fig. 2: Endothelial cells from bovine corpus luteum after 42 days in culture. Longitudinal section of a capillary-like structure with an internal lumen. Transmission electron microscopy, 3,200x.

Quantification of angiogenesis was carried out by evaluation of the stages over the time course.

For verification of reproducibility this method of quantification of angiogenesis was validated. Validation included staging of angiogenesis by different persons and analysis of variability of angiogenesis over time in different culture dishes.

Staging of angiogenesis by different persons

Endothelial cells were incubated with selective medium and observed by phase contrast microscopy over 60 days. After a confluent monolayer was formed, visual fields were chosen randomly and documented twice a week (17 days).

These 68 pictures of cell images were evaluated in random order by two different and independent investigators. Evaluation of cell images depended on the defined stages of angiogenesis *in vitro*. The criterion for the choice of investigators was their comparable knowledge and experience in cultivation and microscopic evaluation of endothelial cells. Each cell image was assigned to a defined stage of angiogenesis. Subsequently, the difference between the stages assigned by the two investigators was calculated for each visual field on each day of investigation. These differences varied in a range between 0 and 1 and, except for two days of investigation, the sum of differences for each day of investigation was ≤ 1.

Additionally, the difference of assigned stages was investigated over the total time (17 days of investigation). For this, the sum of values assigned over 17 days for each visual field and each investigator was calculated (S). Differences in these sums (ΔS) ranged between 0 and 5 (fig. 3) with an arithmetic mean of $\overline{\Delta S}$ = 2,8. Assessment of stages of angiogenesis by different investigators resulted in an inter-person deviation (PD) of 4.1%:

$$PD(4.1\%) - \frac{MVD(2,8)}{MVS(68,1)} x100$$

MVD = Mean value of difference of sum of assigned stages between both investigators

MVS = Mean value of sum of assigned stages by both investigators

Variability of the course of angiogenesis in different culture dishes:

Endothelial cells were seeded in 12 culture wells and cultivated with selective medium. On the first day of investigation six visual fields were chosen randomly in each well and photographed. Three visual fields were chosen in the border area and three visual fields in the centre of the culture well. Thus, investigations were carried out in 72 visual fields, which were defined by coordinates and documented twice a week.

For semiquantitative analysis of angiogenesis in different wells of the culture dishes on each day of investigation, cell images of these 72 visual fields were assigned to one of the defined stages of *in vitro* angiogenesis (1-8).

Documentation of visual fields and thus semiquantitative analysis was carried out over 19 days of investigation.

For each visual field the time dependent course of angiogenesis was investigated by calculating the sum of stages of angiogenesis *in vitro* (S) over all 19 days of investigation. In all 72 visual fields S ranged between 81 and 102, whereas both the minimum (81, slowest course of angiogenesis) and the maximum (102, fasted course of angiogenesis) appeared in the same well.

Homogeneity of the course of angiogenesis in each well was evaluated by calculating variance (s^2) and deviation (s) of the sum of stages (S) of the six visual fields.

Variances of wells ranged between 9.8 and 80.6, its arithmetic mean amounted to $\overline{s^2}$ = 22.1.

This resulted in deviations between 3.1 and 9.0 with a middle deviation of 4.7. Median of variances amounted to 16.6 leading to a deviation of 4.1.

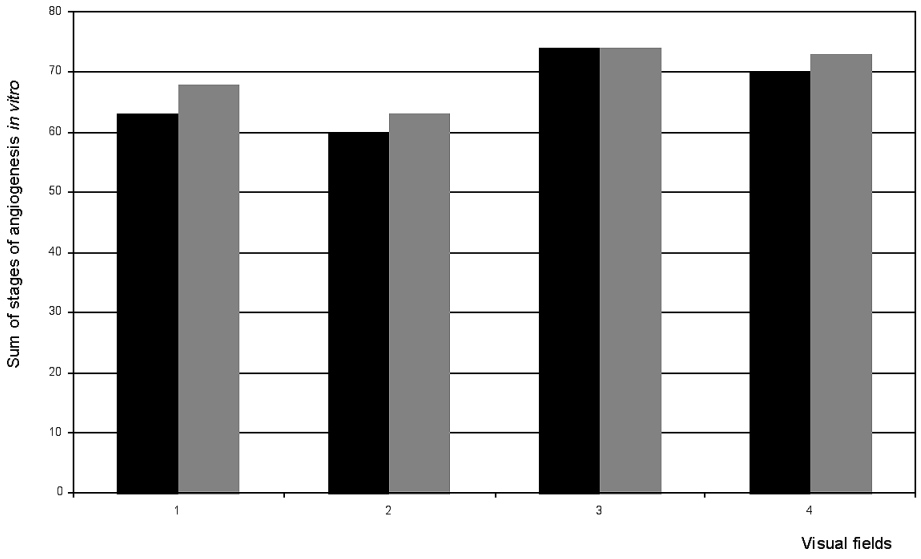


Fig. 3: Staging of angiogenesis by different investigators. Sum of stages of angiogenesis within a period of 60 days in 4 visual fields. Black: Investigator 1; Grey: Investigator 2.

Excepting the well with the highest variance, a middle variance of $s^2 = 16.8$ and a middle deviation of 4.1 were calculated.

Variability of angiogenesis over the course of time in the different culture wells was verified by analysis of variance components (Sachs, 1993) within the culture wells (22.1) and between the culture wells (10.3).

Discussion and conclusions

This report describes the establishment and validation of a new method to quantify angiogenesis *in vitro*. Compared to other *in vivo* and *in vitro* models of angiogenesis this method allows quantification of all phases of angiogenesis by breaking down the angiogenic cascade into quantifiable steps (stages of angiogenesis 1 to 8).

Assessment of stages by different investigators was different in only 19.1% of pictures. These differences always occurred in transition periods between stages and amounted to 1 stage only. Both investigators assigned the stages in the pictures in the correct chronological order starting with stage 1 to stage 8. Thus, results show that quantification of angiogenesis *in vitro* can be carried out by different investigators within the setting of one experiment.

Variability of the course of angiogenesis *in vitro* in different culture dishes was investigated in twelve culture wells. Calculation of variations resulted in deviations of 3.1 to 9.0. However, the deviation of 9.0 diverged widely from the others and was found in the culture well with the highest (102) and lowest (81) sum of stages. When this well was excluded, calculation of the middle variation resulted in $\overline{s^2} = 16.8$ with a middle deviation of $s = 4.1$. Compared to the median of deviation (4.1) of all 12 wells it becomes apparent that the wide divergence in this one well was not representative. Thus, results indicate a justifiable variability of the course of angiogenesis *in vitro* in eleven of twelve culture wells.

Analysis of variance components (Sachs, 1993) showed that the variation between the culture wells was lower than within the culture wells. Thus, the number of culture wells used could be reduced. Further calculations showed that representative quantification of angiogenesis can be carried out in a sample size of four culture wells with the examination of four visual fields per well.

Consequently the *in vitro* model presented allows a viable quantification of angiogenesis *in vitro* based on continuous observation of all stages of the angiogenic cascade. So far this was possible only in a few *in vivo* models of angiogenesis, i.e. the cornea model, the CAM-model and in the dorsal skinfold chamber.

Quantification of angiogenesis in these *in vivo* models however is limited by inflammation induced angiogenesis resulting from insertion of the angiogenic stimulus (Auerbach et al., 2003). Further disadvantages of the *in vivo* models of angiogenesis are the high number of animals required and the high effort of time and costs (Kruger et al., 2001; Donovan et al., 2001). Validation of the *in vitro* method established results in a small sample size which is time and cost efficient.

Depending on surrounding tissue cells, a high variation is found for angiogenesis *in vivo* (Campochiaro and Hackett, 2003). Thus, conclusions drawn from analysing *in vivo* models of angiogenesis are often less significant than assumed. Therefore, one advantage of our and other *in vitro* models of angiogenesis, which mostly are based on monocultures of endothelial cells, is the absence of the influence of other cells and tissues.

In all *in vivo* models quantification of angiogenesis is based on the ascertainment of the number of vessels (Auerbach et al., 2003). However, angiogenesis is defined as the process of vessel generation (Folkman, 2003). Thus, *in vivo* models of angiogenesis only allow quantification of the products of angiogenesis (Bahramsoltani and Plendl, 2004). The new *in vitro* method allows both investigation of molecular mechanisms and their relation to the stages of angiogenesis (Haas, 2005; Ribatti and Ponzoni, 2005; Manson et al., 2005).

Application of angiogenesis stimulators and inhibitors for the treatment of many diseases is in the focus of modern research (Gowda et al., 2005; Markkanen et al., 2005; Huber et al., 2005; Gasparini et al., 2005). Therefore numerous animal experiments are carried out in this promising field.

The newly established and validated *in vitro* method to quantify angiogenesis can be employed either in trial studies of potential angiogenic and angiostatic substances *in vitro*, respectively, or in the investigation of their cellular mechanisms. It may provide an efficient method to replace animal testing, particularly in preclinical screening of new angiogenic or angiostatic substances.

References

Auerbach, R., Lewis, R., Shinners, B. et al. (2003). Angiogenesis assays: a critical overview. *Clin. Chem. 49*, 32-40.

Augustin, H. G. (2003). Translating angiogenesis research into the clinic: the challenges ahead. *Br. J. Radiol. 76*, S3-S10.

Bahramsoltani, M. and Plendl, J. (2004). Ein neues in vitro Modell zur Quantifizierung der Angiogenese. *ALTEX 21*, 227-244.

Bahramsoltani, M. (2004). *Quantifizierung der Angiogenese und Antiangiogenese in vitro*. Berlin: Mensch & Buch Verlag.

Banning, M. (2005). The principles of inflammation in the development of rheumatoid arthritis. *Br. J. Nurs. 14*, 277-283.

Caldwell, R. B., Bartoli, M., Behzadian, M. A. et al. (2005). Vascular endothelial growth factor and diabetic retinopathy: role of oxidative stress. *Curr. Drug Targets 6*, 511-524.

Campochiaro, P. A. and Hackett, S. F. (2003). Ocular neovascularization: a valuable model system. *Oncogene 22*, 6537-6548.

Donovan, D., Brown, N. J., Bishop, E. T. and Lewis, C. E. (2001). Comparison of three in vitro human 'angiogenesis' assays with capillaries formed in vivo. *Angiogenesis 4*, 113-121.

Folkman, J. (2003). Fundamental concepts of the angiogenic process. *Curr. Mol. Med. 3*, 643-651.

Folkman, J. (1995). Angiogenesis in cancer, vascular, rheumatoid and other disease. *Nature Med. 1*, 27-32.

Gasparini, G., Longo, R., Fanelli, M. and Teicher, B. A. (2005). Combination of antiangiogenic therapy with other anticancer therapies: results, challenges, and open questions. *J. Clin. Oncol. 23*, 1295-1311.

Gowda, R. M., Khan, I. A., Punukollu, G. et al. (2005). Treatment of refractory angina pectoris. *Int. J. Cardiol. 101*, 1-7.

Haas, T. L. (2005). Endothelial cell regulation of matrix metalloproteinases. *Can. J. Physiol. Pharmacol. 83*, 1-7.

Hasan, J., Shnyder, S. D., Bibby, M. et al. (2004). Quantitative angiogenesis assays in vivo – a review. *Angiogenesis 7*, 1-16.

Hoying, J. B. and Williams, S. K. (1996). Measurement of endothelial cell migration using an improvement linear migration assay. *Microcirculation 3*, 167-174.

Huber, P. E., Bischof, M. and Jenne, J. (2005). Trimodal cancer treatment: beneficial effects of combined antiangiogenesis, radiation, and chemotherapy. *Cancer Res. 65*, 3643-3655.

Kruger, E. A., Duray, P. H., Price, D. K. et al. (2001). Approaches to preclinical screenings of antiangiogenic agents. *Sem. Oncology 28*, 570-576.

Manson, M. M., Farmer, P. B., Gescher, A. and Steward, W. P. (2005). Innovative agents in cancer prevention. *Recent Results Cancer Res. 166*, 257-275.

Markkanen, J. E., Rissanen, T. T., Kivela, A. and Yla-Herttuala, S. (2005). Growth factor-induced therapeutic angiogenesis and arteriogenesis in the heart – gene therapy. *Cardiovasc. Res. 65*, 656-664.

Meyer, G. T., Matthias, L. J., Noack, L., et al. (1997). Lumen formation during angiogenesis in vitro involves phagocytic activity, formation and secretion of vacuoles, cell death, and capillary tube remodelling by different populations of endothelial cells. *Anat. Rec. 249*, 327-340.

Montesano, R., Orci, L. and Vassalli, P. (1983). In vitro rapid organization of endothelial cells into capillary-like networks is promoted by collagen matrices. *J. Cell Biol. 97*, 1648-1652.

Nehls, V. and Drenckhahn, D. (1995). A novel, microcarrier-based in vitro assay for rapid and reliable quantification of three-dimensional cell migration and angiogenesis. *Microvasc. Res. 50*, 311-322.

Peters, K., Schmidt, H. Unger, R. E. et al. (2002). Software-supported image quantification of angiogenesis in an in vitro culture system: application to studies of biocompatibility. *Biomaterials 23*, 3413-3419.

Plendl, J., Gilligan, B. J., Wang, S.-J. et al. (2002a). Primitive endothelial cell lines from the porcine embryonic yolk sac. *In Vitro Cell Dev. Anim. 38*, 334-342.

Plendl, J., Hirschberg, R. and Hünigen, H. (2002b). Mechanismen der vaskulären Entwicklung und Regression. *Tierärztl. Praxis 30*, 243-253.

Plendl, J. (2000). Angiogenesis and vascular regression in the ovary. *Anat. Histol. Embryol. 29*, 257-266.

Plendl, J. (1997). *Das Gefäßendothel: organ- und entwicklungsphasenspezifische Heterogenität im in vitro Modell*. München: Tierärztliche Fakultät der Ludwig-Maximillians-Universität.

Plendl, J., Neumüller, C. and Sinowatz, F. (1996). Differences of microvascular endothelium in the bovine corpus luteum of pregnancy and the corpus luteum of the estrous cycle. *Bio. Cell. 87*, 179-188.

Pozzi, A. Macias-Perez, I., Abair, T. et al. (2005). Characterization of 5,6- and 8,9-Epoxyeicosatrienoic acids (5,6- and 8,9 EET) as potent in vivo angiogenic lipids. *J. Biol. Chem. 280*, 27138-27146.

Ribatti, D. and Ponzoni, M. (2005) Antiangiogenic strategies in neuroblastoma. *Cancer Treat. Rev. 31*, 27-34.

Rookmaaker, M. B., Verhaar, M. C., Loomans, C. J. et al. (2005). CD34+ cells home, proliferate, and participate in capillary formation, and in combination with CD34-cells enhance tube formation in a 3-dimensional matrix. *Arterioscler. Thromb. Vasc. Biol. 29*, 1843-1850.

Sachs, L. (1993). *Statistische Methoden*. Berlin: Springer-Verlag.

Wang, H., Riha, G. M., Yan, S. et al. (2005). Shear stress induces endothelial differentiation from a murine embryonic mesenchymal progenitor cell line. *Arterioscler. Thromb. Vasc. Biol. 25*, 1817-1823.

Acknowledgement

We thank the Zentralstelle zur Erfassung und Bewertung von Ersatz- und Ergänzungsmethoden zum Tierversuch (ZEBET) of the Bundesinstitut für Risikobewertung (BfR), Berlin, Germany, for financial support.

Correspondence to

Dr. Mahtab Bahramsoltani
Institut für Veterinär-Anatomie
Freie Universität Berlin
Koserstr. 20
14195 Berlin, Germany
e-mail: bahramsoltani.mahtab@vetmed.fu-berlin.de

The Importance of Good Cell Culture Practice (GCCP)

Michael Balls[1], Sandra Coecke[2], Gerard Bowe[2], John Davis[3], Gerhard Gstraunthaler[4], Thomas Hartung[2], Robert Hay[5], Otto-Wilhelm Merten[6], Anna Price[2], Leonard Schechtman[7], Glyn Stacey[8] and William Stokes[9]

[1]FRAME, Nottingham, UK; [2]European Centre for the Validation of Alternative Methods (ECVAM), Ispra (VA), Italy; [3]Research and Development Department, Bio-Products Laboratory, Elstree, Herts., UK; [4]Department of Physiology, Innsbruck Medical University, Innsbruck, Austria; [5]ATCC, Manassas, VA, USA; [6]Généthon, Evry, France; [7]National Center for Toxicological Research, Food and Drug Administration, Rockville, MD, USA; [8]Division of Cell Biology and UK Stem Cell Bank, National Institute for Biological Standards and Control, Blanche Lane, Potters Bar, Herts., UK; [9]National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods Environmental Toxicology Program, National Institute of Environmental Health Sciences, Research Triangle Park, USA

Summary

Following a suggestion made at the 3rd World Congress on Alternatives and Animal Use in the Life Sciences (1999) and the subsequent publication in 2002 of outline guidelines on Good Cell Culture Practice (GCCP), a second ECVAM Task Force was convened, with a broader range of expertise in cell and tissue culture, in order to produce an updated and more-detailed GCCP guidance document for practical use in the laboratory. This GCCP Guidance, which has been published in ATLA in 2005 and is being made freely available, is based on six operational principles, which are briefly summarised in this paper.

Keywords: best practice, cell culture, guidance

Introduction

The use of *in vitro* systems is likely to expand dramatically in the future, not only in basic research, but also: to meet regulatory requirements for chemicals and products of various kinds; in the manufacture of various biological products; in medical diagnostics; and in therapeutic applications such as tissue engineering, and cell and gene therapy. Further significant developments are certain to result from the use of *in vitro* systems for high throughput screening in pharmacology and toxicology; the human genome project; the emerging fields of genomics, proteomics and metabonomics; and the use of biomarkers of disease, susceptibility, exposure and effect. Bearing this in mind, and because the maintenance of high standards is fundamental to all good scientific practice, and is essential for maximising the reproducibility, reliability, credibility, acceptance and proper application of any results produced, it was proposed in 1999, at a workshop held during the 3rd World Congress on Alternatives and Animal Use in the Life Sciences (Hartung and Gstraunthaler, 2000), that guidelines should be developed to define minimum standards in cell and tissue culture, to be called Good Cell Culture Practice (GCCP). This proposal was endorsed by the Bologna Congress participants as a whole (Hartung et al., 2000).

Later that year, an ECVAM Task Force on GCCP was established, which produced outline guidance in 2002 (Hartung et al., 2002). A second ECVAM Task Force was convened in 2003, with a broader range of expertise in cell and tissue culture, in order to produce a more-detailed GCCP guidance document, which could be of practical use in the laboratory. Its report, "Guidance on Good Cell Culture Practice", was published in *ATLA* in 2005 (Coecke et al., 2005), and is being made freely available by various routes.

The aim of this Guidance is to encourage the maintenance of best practice and to reduce uncertainty in the development and application of animal and human cell and tissue culture procedures and products, by encouraging greater international harmonisation, rationalisation and standardisation of laboratory practices, quality control systems, safety procedures, recording and reporting, and compliance with laws, regulations and ethical principles. The scope of the document has deliberately been broadly defined, to include systems based on cells and tissues obtained from humans and animals, and issues related to the characterisation and maintenance of essential characteristics, as well as quality assurance, recording and reporting, safety, education and training, and ethics.

The Guidance is intended to promote high standards in all aspects of the use of cells and tissues *in vitro*, and to complement, but not to replace, any existing guidance, guidelines or regulations, including the guidance on Good Laboratory Practice and Good Manufacturing Practice.

Outline of the GCCP Guidance

The Guidance is based on six operational principles, the basis of which will be briefly summarised.

Establishment and maintenance of a sufficient understanding of the *in vitro* system and of the relevant factors which could affect it

This principle is concerned with the essentials of assuring reliability and accuracy when using cell and tissue-based systems, namely: authenticity (including the identity of the system, its provenance, and confirmation of genotypic and/or phenotypic characteristics); purity (freedom from contamination); stability; and functional integrity. The standardisation of *in vitro* systems begins with the original animal or human donor, then the cells or tissues derived and their subsequent manipulation, maintenance and preservation.

The Guidance therefore deals with: isolated organs or tissues, primary cultures and early passage cultures, and cell lines; *in vitro* culture conditions (media, sera, additives, antibiotics); handling and maintenance (temperature, atmosphere, pH); cell detachment and subculture; cryopreservation; microbial and viral contamination; and cellular cross-contamination.

Assurance of the quality of all materials and methods, maintain the integrity, validity, and reproducibility of any work conducted

This principle is focused on quality issues related to ensuring the consistency, traceability, and reproducibility of *in vitro* work with cells and tissues. Each laboratory should have designated persons responsible for overseeing the quality assurance of: cells and tissues; media and other materials; methods, protocols and SOPs; equipment and its maintenance; recording procedures; and expression of results. Some of these matters are the responsibility of suppliers, rather than end users, and this may involve international standards, such as European Norms, ISO standards or pharmacopoeial protocols.

In addition to the procedures for cells and tissues and other materials, appropriate procedures are needed for items such as low temperature storage systems, incubators, gas cylinders and pressure regulators, laminar air flow and safety cabinets, automatic pipettes and pipettors, sterilisation ovens and autoclaves, and analytical and production equipment.

Documentation of the information necessary to track the materials and methods used, to permit the repetition of the work, and to enable the target audience to understand and evaluate the work

As in any practical science, clear documentation of the systems used and procedures followed is mandatory, in order to permit the traceability, interpretation and repetition of the work. This documentation should include details of: the objective of the work; the rationale of the approach; the materials, equipment, protocols and SOPs used; the origin and characterisation of cells and tissues used; cell and tissue preservation and storage; and the laboratory records, including results, raw data and quality control data.

The format of a report on the work will depend on the target audience, e.g. in-house personnel, a client or sponsor, a regulatory body, the scientific community, or the general public. The person(s) responsible for the report should be identified. Where appropriate, the report should be formally authorised for its intended purpose. A high-quality scientific report should cover the objective of the work, the protocols and SOPs used, planning and experimental design, the execution of the study, data collection and analysis, and a discussion of the outcome.

It should also be made clear that the whole study was established and performed in accordance with any relevant standards, regulations, statutes, guidelines or guidance documents, and safety and quality assurance procedures. A statement of compliance with the GCCP principles should also be included.

The Guidance suggests that, when a report on cell or tissue culture work is submitted for publication, journal editors should require a minimum set of information, which should cover the origins of the cells or tissues, their characterisation, maintenance and handling, and the procedures used. Examples are given to illustrate how this requirement might be applied in the case of a cell line (mouse 3T3 cells) and in the case of primary/early passage human cell cultures.

Establishment and maintenance of adequate measures to protect individuals and the environment from any potential hazards

National and local laws govern safety in the workplace in most countries. Many countries also issue guidelines on occupational health and laboratory safety, and individual laboratories may also have rules which reflect local circumstances.

The Guidance on safety in the cell culture laboratory in no way replaces these laws and regulations, but highlights issues specific to the *in vitro* culture of animal and human cells and tissues. In many countries, each laboratory is required to appoint a "biological safety officer", and this individual should be engaged in the safety evaluation of any cell culture procedures.

In the laboratory, where hazards may be complex and their evaluation requires specialist knowledge, risk assessment should be performed in a structured way. Furthermore, the results of such risk assessments should be recorded, not only to confirm that they have been carried out and appropriate action taken, but also to act as a reference document for individuals performing the tasks assessed. These assessments should be reviewed at regular intervals, to take into account any changes in local practice, national or international regulations, or increases in scientific knowledge.

Particular attention should be paid to risks which may be specific to, or more significant in, certain groups of workers. For example, where there is the possibility that women of reproductive age would be at greater risk from the effects of certain chemicals, such as teratogens or biological agents, and where persons have a diminished immune response.

The safety conditions highlighted relate not only to the safety of individual cell and tissue culture workers, but also to that of their colleagues, the general public and the environment.

All personnel must be made aware of the potential hazards associated with their work, and must be trained in the designated safety procedures, as well as in the appropriate use of the safety

equipment required (including personal protective equipment) and the appropriate handling of spillages.

The risk assessment should include the following: facilities (e.g. laboratories, offices, storage and sanitation), security, especially where special security precautions required, health and safety of staff, laboratory equipment, infectious/biohazardous materials, chemicals and radioactive substances, hazard prevention, transport, and waste disposal.

Compliance with relevant laws and regulations, and with ethical principles

From an ethical and legal point of view, it is desirable that high standards for cell and tissue culture should be established and maintained worldwide, so that accountability, safety and ethical acceptability can be universally guaranteed, insofar as that it is reasonably practicable. The ethical and associated legal issues raised are extremely complex and beyond the scope of the Guidance. However, all concerned should maintain a sufficient level of awareness of the ethical issues related to cell and tissue culture work, of public opinion, and of the relevant legislation at the national and international levels.

At present, there are no ethical guidelines relating specifically to general cell culture practices, but various guidelines, regulations and laws are in place for dealing with cells and tissues of specific origin and/or use.

From a general perspective, diligence in legal and ethical matters leads to data of higher value, since it can hep to avoid waste of effort and encourage greater confidence in the outcome of the study, to the benefit of all concerned, including the general public. The more specific considerations include the ethical implications of: using material of animal origin (in the light of the Three Rs, and including concerns about endangered species, monoclonal antibodies, the use of fetal cells or tissues, the collection of serum [especially fetal bovine serum], and the pretreatment of animals [e.g. to induce CYP450 enzymes]); using material of human origin (origin, use of fetal material, consent, confidentiality, ownership, patents, safety); transplantation, regenerative medicine, and stem cell research and therapy; and genetic manipulation (including the creation, storage, transport, use and disposal of genetically engineered cells).

Provision of relevant and adequate education and training for all personnel, to promote high quality work and safety

The range of applications for cell culture is expanding rapidly and involves an ever-broadening range of technical manipulations for use in basic and applied science, manufacturing, diagnosis, and efficacy and safety testing procedures, as well as for providing therapeutic materials.

The competence of staff to perform their duties is central to ensuring that work is performed according to the standards of the organisation in relation to its scientific, legal and safety requirements and obligations. This requires education and training, as well as the regular monitoring of performance.

A good basic education should be given in the nature and purposes of cell and tissue culture, which is an essential basis for any future training programme. The basic principles of *in vitro* work, aseptic technique, cell and tissue handling, quality assurance, and ethics, should be included. It is also important that those working with material of animal or human origin should have a sufficient understanding of any additional laws or regulations that will apply to them.

Much of the training required may best be given on a one-to-one basis in the laboratory. However, there are a number of principles that can be covered in organised courses that may involve participants from more than one laboratory.

Training should be seen as an ongoing process for improving and developing practical skills and maintaining competence. Given its critical importance, there should be a formally documented training programme for all members of staff, including training records and regular reviews of training needs. To ensure the quality of work in the long term, it is also important to link training with personal development programmes for technical and scientific staff, to ensure they are progressively trained and educated in line with changing laboratory activities and demands.

When new staff join a laboratory, their skills and experience should be assessed, and the need for specific further training in relation to their new jobs should be identified.

Training can be provided in-house by experienced members of staff and/or visiting experts, via accredited on-line programmes and/or through attendance at external courses. For certain applications, including product manufacture and testing, and the processing of cells and tissues for clinical use, training must be formally recorded and reviewed.

Concluding remarks

An electronic version of the Guidance is freely available for use in any ways which serve the interests of GCCP. Its authors hope that it will be widely used to establish and maintain best practice in all aspects of cell and tissue culture work, not only by those who are responsible for such work, but also by those who publish its outcomes or who make important policy decisions based on the information it provides.

These proceedings also contain further comments on the need for GCCP guidelines, by one of the members of the ECVAM Task Force, illustrated by some specific examples (Gstraunthaler, 2006).

References

Coecke, S., Balls, M., Bowe, G. et al. (2005). Guidance on Good Cell Culture Practice. A report of the Second ECVAM Task Force on Good Cell Culture Practice. *ATLA 33*, 261-287.

Gstraunthaler, G. (2006). The need for specific Good Cell Culture Practice guidelines in the cell culture laboratory. These Proceedings.

Hartung, T., Balls, M., Bardouille, C. et al. (2002). Good Cell Culture Practice. *ATLA 30*, 407-414.

Hartung, T. and Gstraunthaler, G. (2000). The standardisation of cell culture procedures. In M. Balls, A-M. van Zeller and M. E. Halder (eds.), *Progress in the Reduction, Refinement and Replacement of Animal Experimentation* (1655-1658). Amsterdam: Elsevier.

Hartung, T., Gstraunthaler, G. and Balls, M. (2000). Bologna Statement on Good Cell Culture Practice (GCCP). *ALTEX 17*, 38-39.

Correspondence to
Prof. Michael Balls
FRAME
Russell & Burch House
96-98 North Sherwood Street
Nottingham NG1 4EE, UK
United Kingdom
e-mail: Michael.Balls@btopenworld.com

Standardisation in Cell and Tissue Culture

The Need for Specific GLP Guidelines in the Cell Culture Laboratory (Good Cell Culture Practice – GCCP)

Gerhard Gstraunthaler
Department of Physiology and Medical Physics, Innsbruck Medical University, Innsbruck, Austria

Summary
The in vitro cultivation of eukaryotic cells has become a powerful technique in basic cell and molecular biological research, applied biotechnology, and in vitro alternatives. Before cell culture could be carried out successfully, two problems had to be overcome: (1) Populations of cells had to be established from single cells; and (2) these populations had to be maintained for many generations. In a successful propagation of cells in vitro, cells from various tissues should grow and proliferate under appropriate culture conditions, while preserving highly differentiated functions, which closely resemble their ancestor cells in vivo. For this purpose, the cell environment in vivo has to be mimicked in vitro. Thus, a number of selected culture conditions and methodological protocols have to be applied, which are handled quite differently in different tissue culture laboratories. Therefore, for good laboratory practice in the cell culture laboratory, well-defined and precisely described culture protocols are mandatory to ensure optimal and reproducible culture conditions, and to enable interlaboratory comparability of data and scientific results obtained with cultured cells.

Keywords: good cell culture practice (GCCP), cell culture, culture media, extracellular matrix, permeable culture supports

Introduction

Cell and tissue culture is defined as (1) the maintenance or cultivation of human or animal cells *in vitro*, including the culture of single cells (cell culture), or (2) the maintenance or growth of tissues *in vitro*, in a way that may allow differentiation and preservation of their architecture and/or function (tissue culture) (Schaeffer, 1990).

In the last decades, cell cultures have become indispensable tools in basic research and applied biomedical sciences. A number of applications of human and animal cell cultures were developed and introduced successfully: (i) to study the normal physiology and biochemistry of cells and their transformed counterparts (i.e. cancer cells), (ii) to test the effect of compounds on specific cell types, (iii) to produce artificial tissue (e.g. tissue engineering), and (iv) to synthesise valuable products (biologicals) in large-scale bioreactors. This increased use of *in vitro* methodologies is indeed a beneficial development in terms of the 3R concept of Russel and Burch (Balls et al., 1995), decreasing the number of experimental animals.

The major advantage of using cell culture is the consistency and reproducibility of results that can be obtained using a batch of cells of a specific type, or, preferably, a homogenous clonal cell population maintained under Good Laboratory Practice (GLP)-based culture protocols.

Human and animal cell culture is rooted in two fundamental concepts in biology. On the one hand on the cellular concept of Schwann and Schleiden in the early 19th century, postulating the cell as the fundamental unit of life, and on the other hand on the concept of homeostasis, the constant maintenance of the internal milieu within tissues, organs, and organisms (McKeehan et al., 1990).

For successful growth and maintenance of human or animal cells *in vitro*, either primary cultures or continuous cell lines, appropriate culture conditions are required that optimally mimic the physiological conditions (internal milieu) *in vivo et situ* (Gstraunthaler, 2003). Thus, the microenvironment of a cell has to be established *in vitro* and be provided by the culture system: (i) temperature, (ii) extracellular ion milieu and osmolality, (iii) extracellular pH and buffering, (iv) basal supply with essential nutrients and oxygen, (v) supplementation with growth factors and hormones, (vi) culture substrates and growth supports, and (vii) disposal of metabolic end products.

As can be expected, a number of variables can be introduced into cell and tissue culture: (i) the supplementation of culture medium with growth factors or differentiation factors, (ii) the use of appropriate culture substrates and/or specific extracellular matrix components, (iii) the subcultivation intervals and seeding densities, (iv) the feeding cycles, and (v) stationary cultures *versus* dynamic medium supply in perfusion reactors. All in all, a number of tissue culture parameters have to be defined and coordinated. However, despite the widespread use and broad applications of cell and tissue cultures, a significant number of basic questions and methodological protocols are still unsolved and are handled in different ways by cell culturists.

In this brief overview, selected examples will be presented on how culture medium composition, medium volume, feeding cycle, serum supplementation, or use of extracellular matrix components will influence growth of cultured cells and the expression of differentiated functions, which represents a seri-

ous impact on the credibility, reliability, reproducibility, and comparability of *in vitro* alternatives.

Culture medium

The culture medium is one of the most important single factors in cell and tissue culture. The culture medium must supply all essential nutrients for cell metabolism, growth and proliferation (Barnes et al., 1987; Butler and Jenkins, 1989; Ham and McKeehan, 1979). These include biosynthetic precursors for cell anabolism, catabolic substrates for energy metabolism, vitamins and trace elements whose function is primarily catalytic, and bulk inorganic ions (electrolytes) whose functions are both catalytic and physiological, e.g. to maintain culture medium pH and osmolality within acceptable limits.

Normal osmolality of human extracellular fluids (plasma and interstitium) is about 290 mosmol/kg. Thus, it is reasonable to assume that this is the optimum osmolality for human and mammalian cells *in vitro*. Certain variations have to be taken into account for other species (e.g. amphibia, insects).

In the renal medulla, however, an interstitial osmolality of up to 1200 mosmol/kg is generated by the countercurrent system of the kidney, providing the basis for the urinary concentrating mechanism in antidiuresis for salt, water, and volume homeostasis. When cultured renal cells of distal tubule, collecting duct, or papillary origin, like MDCK and PAP-HT25 cells, were adapted to hypertonic culture conditions by adding NaCl, raffinose, or urea to the culture medium, the cells responded with the accumulation of organic osmolytes as seen in renal medullary and papillary cells *in vivo* (Handler and Kwon, 1993). This strategy of "hypertonic stress" in culture enabled renal physiologists to study the cellular and molecular mechanisms of renal medullary osmoadaptation *in vitro* (Burg, 1995; Handler and Kwon, 2001).

In the nephron, proximal tubular epithelial cells are capable of gluconeogenesis from C3-precursors, like lactate, pyruvate or alanine. The application of "metabolic stress" by culturing LLC-PK1 and OK proximal tubule-like renal cell lines under glucose-free culture conditions resulted in the selection of gluconeogenic substrains, termed LLC-PK_1-$FBPase^+$ and OK_{GNG+}, respectively (Gstraunthaler and Handler, 1987; Gstraunthaler et al., 1993). Both cell lines express the key gluconeogenic enzymes phosphoenolpyruvate carboxykinase (PEPCK) and fructose-1,6-bisphosphatase (FBPase). Furthermore, LLC-PK_1-$FBPase^+$ cells respond to metabolic acidosis, i.e. a decrease in culture medium pH from 7.4 to 6.9, with adaptive increases in PEPCK activities and enhanced glutamine metabolism (Gstraunthaler et al., 2000; Holcomb et al., 1995), making it a valuable *in vitro* model to study renal acid-base adaptation and pH-mediated gene expression at the cellular and molecular level (Curthoys and Gstraunthaler, 2001).

Oxygenation and culture medium volumes

Human and animal cells are characterised by oxidative metabolism. Thus, a sufficient supply of cultured cells with oxygen is critical to satisfy cell metabolism. In small cultures (e.g. T-flasks, culture dishes), the oxygen demand can be met by gas diffusion from the headspace through the culture surface. Under normal culture conditions, the availability of oxygen to cells growing under 3-6 mm of medium is adequate.

When renal proximal tubular cells are brought into tissue culture, they revert from oxidative metabolism and gluconeogenesis to high rates of glycolysis. Among the factors possibly responsible for this metabolic conversion, limited oxygen availability and/or substrate supply are discussed. In order to study the role of these factors in long-term cultures, the impact of growth conditions, culture medium volume, and glucose content on carbohydrate metabolism of the continuous renal cell lines LLC-PK_1 (porcine kidney) and OK (opossum kidney) was investigated (Gstraunthaler et al., 1999).

The impact of both culture medium volume and glucose content was determined by overlaying confluent monolayer cultures of LLC-PK_1 and OK cells (i) with increasing volumes of culture medium, thereby increasing the amount of glucose and increasing the diffusion distance for oxygen, and (ii) with increasing culture medium volumes at constant absolute amounts of glucose by adding glucose-free medium, in order to increase volume and thus decrease oxygen supply at a constant glucose supply. Alternatively, and in order to improve cell oxygenation, LLC-PK_1 cells were also cultured in roller bottles. Cell carbohydrate metabolism was assessed by measuring rates of glucose consumption and lactate production, respectively, and by determination of specific activities of the key glycolytic enzymes hexokinase (HK), phosphofructokinase (PFK), pyruvate kinase (PK), and lactate dehydrogenase (LDH). Mitochondrial phosphate-dependent glutaminase (PDG) was assayed as marker enzyme for oxidative metabolism of glutamine.

In LLC-PK_1 and OK cells, rates of glucose consumption were independent of the initial glucose concentrations and/or the culture medium volumes used. Glucose was quantitatively converted to lactate, which accumulated in a 1:2 molar ratio. Lactate in the culture medium reached a maximum content after 24 h and was re-utilised by the cell lines thereafter. Interestingly, the rates of lactate re-uptake strictly depended on culture medium volume, indicating a volume-induced stimulation of oxidative lactate metabolism. Marked changes were found for the specific activities of glycolytic enzymes. In LLC-PK_1 cells, increased glucose supply caused increases in HK, PFK, PK and LDH activities, which were superimposed on the stimulatory effects of increased medium volumes. Enzyme activity showed a biphasic response, indicating that both glucose supply and culture medium volume covering the cell monolayer, and thus oxygen availability, are factors determining glycolytic rates of LLC-PK_1 renal cells. As expected, under conditions of enhanced oxygenation of LLC-PK_1 cells in roller bottle culture, glycolytic enzyme activities decreased, whereas PDG activity increased, which was paralleled by increased rates of ammonia generation. Thus, changes in nutrient supply and oxygenation of renal epithelial cell cultures by altered culture medium volumes dramatically influence metabolic rates and levels of enzyme activities (Gstraunthaler et al., 1999).

Culture substrate and cell adhesion

Most cultured cells, primary cultures as well as continuous cell lines, are substrate- or anchorage-dependent, which means that the cells require attachment to a given surface (culture substrate) in order to survive, proliferate, and express their differentiated functions (Ruoslahti and Öbrink, 1996; Balda and Matter, 2003). In routine cell culture this is mostly accomplished by the use of surface-charged, hydrophilic polystyrene culture plastic ware, that allows crosslinkage of cells via glycoproteins and/or divalent cations (Ca^{2+}, Mg^{2+}). In special cases, culture vessel surfaces can be coated with components of the extracellular matrix, like collagen type IV, laminin or fibronectin to allow attachment via specific cell receptors. Also, collagen type I, gelatin or poly-D-lysine are used as surface coatings to improve adhesion and differentiation of cultured cells (Genestie et al., 1997; Gumbiner, 1996; Kleinman et al., 1987).

When MDCK renal epithelial cells were cultured on a complex extracellular matrix, i.e. basement membrane gel extracted from Engelbreth-Holm swarm tumour, the cells differentiated into tall epithelia, columnar in shape, with an ultrastructure characteristic for fluid-transporting renal epithelia (Zuk et al., 1989).

Culture dish *vs.* permeable filter supports

Traditionally, cultured cells are grown on impermeable culture substrates. Under these conditions, cultured renal epithelial cells form monolayers of highly differentiated and polarised cells. The cells grow with their basolateral surface on the bottom of the culture dish and thus, the apical surface faces the culture medium. Under these conditions, cultured epithelial cells can form domes, i.e. fluid-filled blisters generated by the transepithelial vectorial transport of fluid and solutes trapped between the cell layer and the water-impermeable culture dish (Gstraunthaler, 1988). However, access to growth medium on the basolateral side, at which nutrient exchange normally occurs, is restricted. In contrast, permeable, microporous membrane substrates provide independent access to both the apical and the basolateral side of the cultured epithelium, and thus offer a more physiologically relevant environment (Handler et al., 1984).

The apical and the basolateral fluid compartments are separated by the cultured epithelium, which allows determination of electrophysiological and transport parameters in Ussing-type chamber devices. It has been shown in a number of studies that the degree of differentiation considerably increases when the epithelia are cultured on permeable surfaces. Also, remarkable differences in cell shape and cell density were observed in filter-grown epithelia (Genestie et al., 1997; Gstraunthaler et al., 2000). A6 cells, for example, derived from the kidney of *Xenopus laevis*, respond to vasopressin with an increased rate of sodium transport only when cultured on permeable filter supports. Functional assays revealed that A6 cells do not express vasopressin receptors until an ordered epithelium has been formed on permeable culture supports (Lang et al., 1986).

This technique of culturing transporting epithelia represents the prerequisite for studying epithelial dysfunction in *in vitro* nephrotoxicity testing (Gstraunthaler et al., 1990; Steinmassl et al., 1995).

In conclusion, a number of tissue culture parameters substantially influence the expression of specific morphological features and cellular functions and thus the degree of differentiation of cultured cells. A minimum set of standards has to be defined in order to establish reproducibility and interlaboratory comparability of results obtained with *in vitro* cell culture technologies. In analogy to Good Laboratory Practice (GLP), a Good Cell Culture Practice (GCCP) should be implemented in tissue culture work. GCCP guidelines were recently elaborated by two ECVAM Task Forces (Balls et al., these proceedings; Coecke et al., 2005; Hartung et al., 2002).

References

Balda, M. S. and Matter, K. (2003). Epithelial cell adhesion and the regulation of gene expression. *Trends Cell Biol. 13*, 310-318.

Balls, M. et al. (2006). The importance of Good Cell Culture Practice. *ALTEX*, this volume.

Balls, M., Goldberg, A. M., Fentem, J. et al. (1995). The three Rs: the way forward. The report and recommendations of ECVAM Workshop 11. *ATLA 23*, 838-866.

Barnes, D., McKeehan, W. L. and Sato, G..H. (1987). Cellular endocrinology: Integrated physiology in vitro. *In Vitro Cell. Dev. Biol. 23*, 659-662.

Burg, M. B. (1995). Molecular basis of osmotic regulation. *Am. J. Physiol. 268*, F983-F996.

Butler, M. and Jenkins, H. (1989). Nutritional aspects of the growth of animal cells in culture. *J. Biotechnol. 12,* 97-110.

Coecke, S., Balls, M., Bowe, G. et al. (2005). Guidance on Good Cell Culture Practice. A report of the second ECVAM task force on Good Cell Culture Practice. *ATLA 33*, 261-287.

Curthoys, N. P. and Gstraunthaler, G. (2001). Mechanism of increased renal gene expression during metabolic acidosis. *Am. J. Physiol. 281*, F381-F390.

Genestie I., Morin J.-P., Guery C. et al. (1997). Morphological and biochemical characterization of primary culture of rabbit proximal kidney tubule cells grown on collagen-IV coated Millicell-CM. *In Vitro Cell. Dev. Biol. 33*, 692-702.

Gstraunthaler, G. (2003). Alternatives to the use of fetal bovine serum: serum-free cell culture. *ALTEX 20*, 275-281.

Gstraunthaler, G. J. A. (1988). Epithelial cells in tissue culture. *Renal Physiol. Biochem. 11*, 1-42.

Gstraunthaler, G. and Handler, J. S. (1987). Isolation, growth, and characterization of a gluconeogenic strain of renal cells. *Am. J. Physiol. 252*, C232-C238.

Gstraunthaler, G., Steinmassl, D. and Pfaller, W. (1990). Renal cell cultures: a tool for studying tubular function and nephrotoxicity. *Toxicol. Lett. 53*, 1-7.

Gstraunthaler, G., Thurner, B., Weirich-Schwaiger, H. and Pfaller W. (1993). A novel gluconeogenic strain of OK cells with metabolic properties different from gluconeogenic LLC-PK_1 cells. *Cell. Physiol. Biochem. 3*, 78-88.

Gstraunthaler G., Seppi, T. and Pfaller, W. (1999). Impact of culture conditions, culture media volumes, and glucose content on metabolic properties of renal epithelial cell cultures. Are

renal cells in tissue culture hypoxic? *Cell. Physiol. Biochem. 9*, 150-172.

Gstraunthaler, G., Holcomb, T., Feifel, E. et al. (2000). Differential expression and acid-base regulation of glutaminase mRNAs in gluconeogenic LLC-PK_1-$FBPase^+$ cells. *Am. J. Physiol. 278*, F227-F237.

Gumbiner, B. M. (1996). Cell adhesion: the molecular basis of tissue architecture and morphogenesis. *Cell 84*, 345-357.

Ham, R. G. and McKeehan, W. L. (1979). Media and growth requirements. *Methods Enzymol. 58*, 44-93.

Handler, J. S. and Kwon, H. M. (1993). Regulation of renal cell organic osmolyte transport by tonicity. *Am. J. Physiol. 265*, C1449-C1455.

Handler, J. S. and Kwon, H. M. (2001). Cell and molecular biology of organic osmolyte accumulation in hypertonic renal cells. *Nephron 87*, 106-110.

Handler, J. S., Preston, A. S. and Steele, R. E. (1984). Factors affecting the differentiation of epithelial transport and responsiveness to hormones. *Fed. Proc. 43*, 2221-2224.

Hartung, T. Balls, M., Bardouille, C. et al. (2002). Good Cell Culture Practice. ECVAM Good Cell Culture task force report 1. *ATLA 30*, 407-414.

Holcomb, T., Curthoys, N. P. and Gstraunthaler, G. (1995). Subcellular localization of PEPCK and metabolism of gluconeogenic substrains of renal cell lines. *Am. J. Physiol. 268*, C449-C457.

Kleinman,, H. K., Luckenbill-Edds, L., Cannon, F. W. and Sephel, G. C. (1987) Use of extracellular matrix components for cell culture. *Anal. Biochem. 166*, 1-13.

Lang, M. A., Muller, J., Preston, A. S. and Handler, J. S. (1986). Complete response to vasopressin requires epithelial organization in A6 cells in culture. *Am. J. Physiol. 250*, C138-C145.

McKeehan, W. L., Barnes, D., Reid, L., et al. (1990). Frontiers in mammalian cell culture. *In Vitro Cell. Dev. Biol. 26*, 9-23.

Ruoslahti, E. and Öbrink, B. (1996). Common principles in cell adhesion. *Exp. Cell Res. 227*, 1-11.

Schaeffer, W. I. (1990). Terminology associated with cell, tissue and organ culture, molecular biology and molecular genetics. *In Vitro Cell. Dev. Biol. 26*, 97-101.

Steinmassl, D., Pfaller, W., Gstraunthaler, G. and Hoffmann, W. (1995). LLC-PK1 epithelia as a model for in vitro assessment of proximal tubular nephrotoxicity. *In Vitro Cell. Dev. Biol. 31*, 94-106.

Zuk, A., Matlin, K. S. and Hay, E. D. (1989). Type I collagen gel induces Madin-Darby Canine Kidney cells to become fusiform in shape and lose apical-basal polarity. *J. Cell Biol. 108*, 903-919.

Correspondence to
Gerhard Gstraunthaler, Ph.D.
Innsbruck Medical University
Department of Physiology and Medical Physics
Fritz-Pregl-Strasse 3
6020 Innsbruck, Austria
e-mail: gerhard.gstraunthaler@i-med.ac.at

Ensuring Quality of *In Vitro* Alternative Test Methods

Amy S. Rispin[1] *and Katherine Stitzel*[2]
[1]U.S. Environmental Protection Agency, Office of Pesticide Programs (7506C), Washington, D.C., U.S.A.
[2]Consultant, West Chester, Ohio, U.S.A.

Summary
In vitro and ex vivo methods have been developed or are under development to reduce or replace animal usage in toxicity tests. Consensus is developing in the scientific community on the quality control measures needed for in vitro methods; including appropriate controls, data reporting elements, and benchmarks to be identified in test guidelines so that the potential risks of chemicals can be reviewed and assessed reliably. Consistent with the goal of obtaining scientifically sound test data for hazard and risk assessment of chemicals, changes have been made in current policies and procedures to facilitate the acceptance of data developed using these methods. National and international organisations have developed policies and standards for scientific practice to assure quality in the implementation of in vitro methods. ICCVAM and ECVAM have developed the Performance Standards process to allow proprietary test systems using in vitro/ex vivo methods to be accepted for regulatory use, where Performance Standards include use of reference chemicals, essential test method components and statistical performance results. Additional guidance has been provided for OECD's Good Laboratory Practice principles which will help to ensure that in vitro tests used for regulatory purposes are reproducible, credible and acceptable. Generic test guidelines incorporating Performance Standards are being written to allow acceptance of proprietary test methods by regulatory agencies and to provide assurance that any in vitro system performs over time in a manner that is consistent with the test system as it was originally validated. Future developments should address standardised data reporting elements for special techniques, such as cell and tissue culture or microarrays.

Keywords: in vitro, quality, performance standards, toxicology, testing, cell culture, alternatives, proprietary test methods

Introduction*

OECD members as well as other countries have legislated mandatory submission of credible scientific data for use in assessing the hazards and potential risks of chemicals to humans, wildlife, and the environment. A basic principle in such legislation is that evaluations of the safety of new chemicals must be based on toxicology test data of sufficient quality, rigor, and reproducibility. Although *in vivo* tests in laboratory animals have formed the foundation of hazard and risk assessment, regulatory programmes are now beginning to implement non-animal methods such as cell and tissue culture systems, and high throughput methods such as toxicogenomics and proteomics. These alternatives to current methods often refine, reduce, or replace animal use while providing a comparable or better level of protection of human health or the environment.

Test methods destined for regulatory use evolve in a systematic fashion from research and development through test method design and validation (see fig. 1).

Specific *in vitro/ex vivo* and other non-animal methods have been developed or are currently under development to replace animal tests (ECVAM, 2002) or to allow direct assessment of chemical effects in human cells or tissue components. Areas of alternative assay development are outlined in figure 2. When any new test method is developed, test parameters are standardised so that laboratories can obtain consistent results. Although quality assurance is always an issue, *in vitro* test systems pose different issues regarding their quality and performance than commonly used animal methods (Rispin et al., 2004).

Some *in vitro* assays include a bioconstruct or *ex vivo* component that acts as the target tissue for the toxicological effect of concern. The bioconstruct can be a cellular, non-cellular, or tissue construct. Tissue constructs, often made of materials derived from humans, are designed to model the toxicology of cells or tissues and replicate the *in vivo* responses to chemical exposure. All of the elements of the assay function together for purposes of the assay's use to fulfil regulatory testing requirements. The quality of the bioconstruct and associated reagents must be assured for any assay to be used for regulated studies. Quality control ensures that the results of the assay can be reliably used in hazard and risk assessment and can be compared with data from previous studies within a laboratory and between one laboratory and another. Because these systems can be affected by small changes in method or components, such test systems must be well defined and function reproducibly.

Once a new method is developed, it must be validated (OECD, 2005). Validation involves systematic laboratory studies performed on a set of common reference chemicals to determine the new test's reliability in terms of intra- and inter-laboratory variability, and to assess how well it functions for various chemical classes.

*Much of this material has been previously published in *Regulatory Toxicology and Pharmacology* (Gupta et al., 2005; Rispin et al., 2005 (in press)).

In the United States, the Interagency Coordinating Committee for Validation of Alternative Methods (ICCVAM) provides for review and assessment of the validity of the new toxicology tests including non-animal alternative test systems and proprietary test methods (ICCVAM, 2003). The European Centre for Validation of Alternative Methods (ECVAM) has been established to facilitate development of non-animal tests for the European Union and to assess the reliability and relevance of such tests for European regulatory mandates (ECVAM, 1995). Both ICCVAM and ECVAM are directed to seek alternative tests which reduce, refine, or replace animal testing.

Once a new *in vitro* method is validated and accepted for regulatory use, companies and regulatory authorities making decisions based on the data need assurance that it will continue to perform in a manner consistent with the test system as it was originally validated. Stability of performance of the *in vitro* system must be ensured over time, particularly if there are changes in components of the test system, changes in test system manufacturer(s), and/or changes in the prediction model. Testing laboratories must use good scientific practice, as well as appropriate calibration and standardisation methodology established by the various technical disciplines appropriate to the elements of their assay system.

Special considerations for proprietary test methods

Additional issues arise when the *in vitro* methods are developed, validated and registered by manufacturers for commercial marketing as proprietary test methods. Generally, the process for acceptance of such proprietary test methods calls for special approaches. At OECD, the specific proprietary test system cannot be accepted under the Mutual Acceptance of Data Decision, but this OECD agreement does allow for use of a generic guideline based on the proprietary test method. This allows other companies to enter the marketplace with "me-too" methods; a "me-too" method is mechanistically and functionally similar to a validated method. Regulatory acceptance of these generic methods is based on validation data obtained using the proprietary version. The European Union and OECD use generic guidelines for proprietary test methods, following their validation. If a proprietary test method assay system is identified by a United States agency for use in regulatory testing, the agency cannot mandate use of the specific proprietary test method, but can develop Performance Standards for a generic protocol incorporating the salient features of the proprietary test method such that a competitor may introduce a product that meets the Performance Standards. A United States agency may then state that the proprietary test method meets the Performance Standards. The applicant for approval of a chemical who wishes to use a different proprietary test method may do so as long as documentation is provided to the agency that the manufacturer of the second proprietary test method has assured compliance with the Performance Standards (USEPA, 2003, 2004). OECD has issued three generic guidelines which incorporate elements of Performance Standards for *in vitro* assessment of dermal corrosivity: EpiDerm™/EPISKIN™ – OECD 431; Corrositex® – OECD 435; and Transcutaneous Electrical Resistance Test (TER) – OECD 430.

Once a new proprietary test method is accepted for regulatory use, both regulatory agencies and the users of these test systems need a process to ensure that "me-too" test kits developed according to the generic descriptions produce results similar to those obtained using the system originally validated and accepted.

Quality assurance for regulatory use

International quality control requirements for assays performed to fulfil regulatory requirements are called for under the OECD Mutual Acceptance of Data Decision which sets forth international standards for Good Laboratory Practice (GLP) (OECD, 1981). The principles of GLP have been developed to promote the quality and relevance of test data used for determining the safety of chemicals and chemical products. GLPs apply to non-clinical health and environmental safety studies for the registration of pharmaceuticals, veterinary drugs, pesticides, and food

Fig. 1: Test method evolution

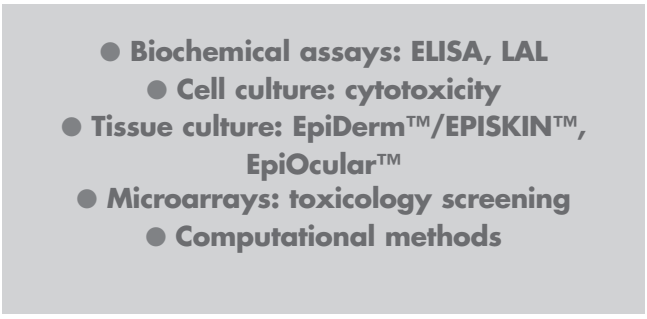


Fig. 2: Areas of alternative assay development

and feed additives, and for the regulation of industrial chemicals, cosmetic products and consumer products.

GLPs are concerned with the organisational process and the conditions under which non-clinical health and environmental safety studies are planned, performed, recorded, archived and reported. Their aim is to assure consistency, traceability and reproducibility of test results. Test facilities are required to follow GLP principles when carrying out studies to be submitted to national authorities for the purpose of assessment of chemicals and other uses related to the protection of man and the environment. GLP principles are accepted throughout OECD countries and world-wide (USEPA, 2002; USFDA, 2003). Within countries, GLPs are implemented by regulatory authorities.

GLP and good science responsibilities of the testing laboratories include using a pre-defined study protocol for the assay, training the technical staff and ensuring that staff skills are maintained, employing controls and calibrated equipment, employing established and defined regulatory acceptability criteria for the test method, and accurately and completely documenting each step in the procedure. GLP principles apply to *in vitro* tests as well as tests using animals. Previously GLP guidelines focused primarily on *in vivo* methods and required significant interpretation to cover *in vitro* methods (Cooper-Hannan et al., 1999).

Policy framework to assure quality of alternative methods

When an alternative test method matures, its regulatory use depends on availability of a comprehensive scientific and policy framework. Due to the efforts of ICCVAM, ECVAM, OECD, scientific societies, and the animal welfare community, a six element framework is now in place internationally to assure quality of *in vitro* alternative methods: OECD Process for Proprietary Test Methods; Performance Standards – ICCVAM; Application of the principles of GLP to computerised systems; Application of the principles of GLP to *in vitro* studies; and Guidance on Good Cell Culture Practice (GCCP) (see fig. 3) (Gupta et al., 2005; Rispin et al., 2005).

OECD consensus document on application of GLP to *in vitro* studies

Some time ago, OECD developed special guidance for application of GLPs to computerised systems (OECD, 1995). More recently, recognising that *in vitro* methods will be more prominent as alternatives or supplements to traditional *in vivo* safety testing, ICCVAM and ECVAM recommended to OECD that it develop special guidance for the application of GLP to *in vitro* studies. The new OECD Advisory Document for *in vitro* studies (OECD, 2004) calls for study directors to document that *in vitro* test methods be validated, or are structurally, functionally and mechanistically similar to a validated reference test method and have comparable performance. In addition, representative batches of test components should be evaluated for their performance.

Performance Standards can provide a framework for application of GLP to *in vitro* assays. (See details in Performance Standards section below.)

The OECD advisory document recognises the growing use of cell/tissue culture and anticipates developments in toxicogenomics, toxicoproteomics, toxicometabonomics and microarrays and is expected to facilitate interpretation of GLP for *in vitro* studies performed for regulatory purposes.

ECVAM task force on good cell culture practice (GCCP)

In 1999, based on development of several promising cell culture test systems as alternatives to animal tests, the Third World Congress proposed that cell culture procedures be standardised to meet a number of testing and regulatory needs (see fig. 4). The first four principles of GCCP apply equally well to microarrays, biochemical assays, or computational methods developed as alternative methods. ECVAM has issued two reports on GCCP (Hartung et al., 2003; Coecke et al., 2005), which specifies procedures to ensure that test systems are free of any contamination or other diseases or conditions at the beginning of the study that might interfere with the outcome of the study and calls for the origin (species/tissue), source, arrival condition and mainte-

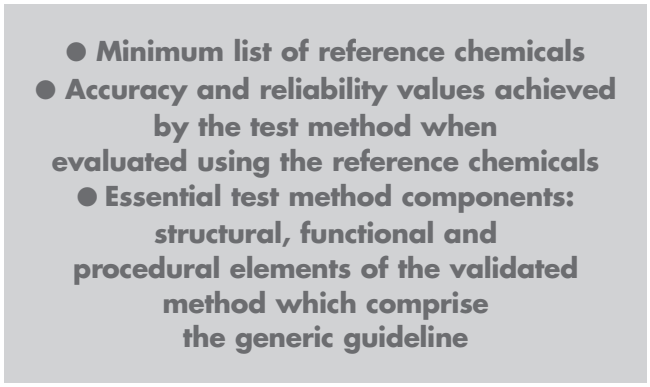


Fig. 3: Framework to assure quality of *in vitro* alternative test methods

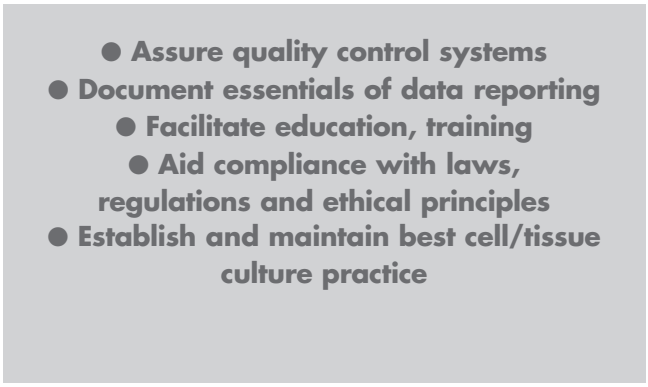


Fig. 4: Good Cell Culture Practice principles

nance requirements to be documented and confirmed at the laboratory on a regular basis. These GCCP reports are intended to set standards for education and training in Europe, to educate the editors of journals to look at the critical cell culture parameters when a paper is submitted, and to help regulatory authorities in the acceptance and interpretation of *in vitro* data.

Scientific societies and other organisations are also developing standardised data reporting elements or quality assurance measures for other promising alternative technologies such as microarrays, software systems, etc.

Performance Standards

The elements of Performance Standards are summarised in figure 5. Although originally planned for regulatory acceptance of Proprietary Test Methods, they provide a comprehensive approach for specifying performance of any alternative test methods. Performance Standards are descriptive and functional and serve to demonstrate that any kit or assay system is mechanistically and functionally equivalent to the test that was initially accepted for regulatory purposes (ICCVAM, 2003). Reference chemicals drawn from the data base used for validation of the original proprietary test method are used to establish functional standards, with "me-too" assays required to perform comparably to the original validated method. Descriptive standards are incorporated into generic assay guidelines and may include consideration of the integrity and structural characteristics of key elements of the assay, such as a tissue or bioconstruct. For example, for cells in culture, descriptive characteristics would involve exact characterisation of the identity and confirmation of the structural integrity of cells.

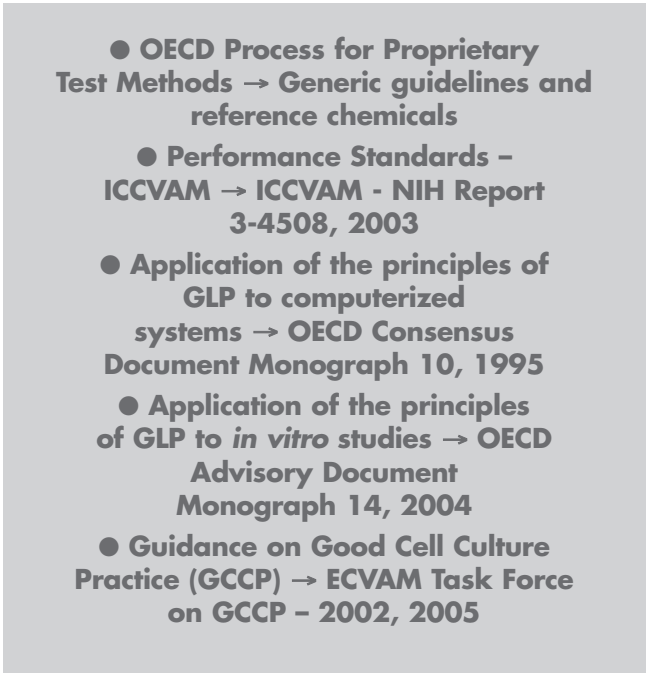


Fig. 5: Performance Standards for proprietary test methods and other validated *in vitro* alternative methods

Much of the activity to develop non-animal methods for toxicological evaluation has been and is expected to be at the behest of commercial sponsors. United States agencies will need Performance Standards for all new proprietary test methods and for other *in vitro* methods as well. Availability of Performance Standards will facilitate the development of new and "me-too" test methods while setting standards for their accuracy and reliability.

Sponsors submitting new test methods to ICCVAM or ECVAM for review may wish to propose Performance Standards for consideration: such Performance Standards will be included in test guidelines approved for use by regulatory agencies. ICCVAM has recently revised its submission guidelines to provide for development of Performance Standards, including essential test method components, reference chemicals, and performance criteria for each test method (ICCVAM, 2003).

The use of additional controls when using *in vitro* methods

The role of experimental controls (positive, negative, or graded) is normally spelled out in agencies' policies for acceptability of test procedures. In recent years, the consistent quality of the test animals as well as concerns for animal welfare have persuaded many authorities to do without concurrent positive controls for most acute tests involving laboratory animals. However, for alternative methods in which animals are not involved, laboratories should use positive as well as negative controls and benchmark chemicals (defined below) as part of every chemical trial using *in vitro* assays. The use of appropriate controls ensures lot-to-lot consistency of the biological or *ex vivo* assay components of proprietary test methods as well as non-proprietary tests and verifies that test systems are calibrated and functioning properly. Such chemicals may be selected from the Performance Standard reference chemicals for the *in vitro* assay.

The assay endpoint value(s) for the negative control should fall within an acceptable range as determined by historical experience with the test system. The negative control test system units, e.g. tissue constructs, are manipulated in parallel with those treated with test and positive control materials. The assay endpoint is measured on these test system units and provides a baseline value for comparison with values from the test system elements treated with test and positive control materials. This baseline may be the maximum viability value or the absence of lesions in a histological evaluation. In viability assays, the negative control value provides the "100% viability" value for determination of the relative viability of the test substance-treated tissues.

The positive control material is used to assess the functional characteristics of the test system and the execution of the assay. The positive control-treated test system units are tested in parallel with the negative control and test material-treated units each time the assay is performed. Positive controls should be chosen so that the test system response falls in the mid range of the pos-

sible responses and both increases and decreases in response against the historical performance of the positive control can be detected. The results from the control assays help to confirm that a valid trial was performed when test data for each unknown chemical are submitted to regulatory agencies. In addition, trend analysis of control trials can be used to detect drift in the assay system.

Benchmark chemicals with well-characterised responses selected from the same structural or functional chemical class as the unknown or test chemical provide assurance that the test is performing correctly for the specific class of compound being tested. A benchmark chemical may be included in each experimental trial along with the positive and negative controls.

The use of concurrent controls is critical to ensure that testing laboratories and regulatory authorities can be confident that *in vitro* results continue to fulfil their intended regulatory function. In addition, properly selected controls can ensure that proprietary test methods are functioning as designed. Toxicologists must be provided with appropriate standardisation or verification approaches, including appropriate controls, to ensure that proprietary assay systems are functioning properly.

Responsibilities for ensuring quality data from *in vitro* methods

Practically speaking, quality control of *in vitro* assay systems must be a shared responsibility of the manufacturer of the proprietary test method or supplier of tissues or tissue constructs, the contract laboratory that uses the assay system, and the company that submits the test for regulatory acceptance. Bioconstructs or cell cultures may have a short shelf life. Proprietary test methods and other *in vitro* systems may be relatively expensive; therefore, the number of replicate systems available for quality control efforts by testing laboratories may be limited by practical considerations. In light of these concerns, the user may be dependent on the manufacturer for many of the basic elements of quality control, including cell or tissue characterisation and functional performance of the assay system. The manufacturer should be expected to provide adequate documentation of quality control testing of representative assay systems for each manufactured batch. In addition, the user must provide a quality control check in the laboratory on a regular basis appropriate to the test system so that the assay materials can be shown to perform as expected after transport and handling. Regulatory authorities will need to develop submission guidelines that are clear and consistent with the realities of the new regulatory test systems.

Summary

Ensuring quality of results from *in vitro* and *ex vivo* test systems in regulatory toxicology testing calls for innovative approaches by all concerned; testing laboratories, regulatory authorities, validation organisations, and the scientific community. In preference to imposing new regulations, existing statutes, directives and implementing regulations are being interpreted to extend their guidance to the many new technologies used to replace or augment traditional animal tests.

When the validity of a new *in vitro* test is reviewed, the reviewing organisation normally considers use of controls and benchmark chemicals and can also recommend Performance Standards. Test system sponsors may wish to propose a set of Performance Standards for use in assuring consistency of the test system response when they prepare the new system for review of its validation status. Performance Standards required for proprietary test methods for use by regulatory agencies can also be part of the foundation for quality control of the proprietary test methods.

With experience, the expectations for quality assurance for non-animal testing will evolve. It will be important for test developers, laboratories performing the tests, toxicologists and regulatory agencies to continue to work together to ensure that guidelines meet the needs of all parties.

References

Coecke, S., Balls, M., Bowe, G. et al. (2005). Second ECVAM Task Force on Good Cell Culture Practice. Guidance on good cell culture practice. A report of the second ECVAM task force on good cell culture practice. *ATLA 33*, 261-287.

Cooper-Hannan, R., Harbell, J. W., Coecke, S. et al. (1999). The principles of Good Laboratory Practice: Application to in vitro toxicology studies. The report and recommendations of ECVAM workshop 37. *ATLA 27*, 539-577.

European Center for Validation of Alternative Methods (ECVAM) (1995). Workshop Report No. 5: Practical Aspects of the Validation of Toxicological Test Procedures. *ATLA 23*, 129-147.

European Center for Validation of Alternative Methods (ECVAM) (2002). *Alternative (Non-animal) Methods for Chemicals Testing: Current Status and Future Prospects*. A. Worth and M. Balls (eds.) ECVAM, European Commission Joint Research Center, 21020, Ispra (VA), Italy.

Gupta, K., Rispin, A., Stitzel, K. et al. (2005). Ensuring quality of in vitro alternative test methods: Issues and answers. *Reg. Tox. Pharm. 43*, 219-328.

Hartung, T, Balls, M., Bardouille, C. et al. (2003). Good Cell Culture Practice: ECVAM Good Cell Culture Practice Task Force Report 1. *ATLA 30*, 407-414.

Interagency Coordinating Committee for Validation of Alternative Methods (ICCVAM) (2003). *ICCVAM Guidelines for the Nomination and Submission of New, Revised, and Alternative Test Methods*. National Institute of Environmental Health Sciences. NIH Publication 03-4508. Research Triangle Park, NC, USA.

Organization for Economic Cooperation and Development (OECD) (1981). Decision of the Council Concerning the Mutual Acceptance of Data in the Assessment of Chemicals. [C(81)30]. Adopted May 12, 1981. Paris, France.

Organization for Economic Cooperation and Development (OECD) (1995). Consensus Document of the Working Group

on Good Laboratory Practice: The Application of the Principles of GLP to Computerized Systems. OECD Series on Principles of Good Laboratory Practice and Compliance Monitoring Number 10. Paris, France.

Organization for Economic Cooperation and Development (OECD) (2004). Advisory Document of the Working Group on Good Laboratory Practice: The Application of the Principles of GLP to in vitro Studies. OECD Series on Principles of Good Laboratory Practice and Compliance Monitoring Number 14 (ENV/JM/MONO(2004)26). Paris, France.

Organization for Economic Cooperation and Development (OECD) (2005). Guidance Document on the Validation and International Acceptance of New or Updated Test Methods for Hazard Assessment. OECD Series on Testing and Assessment Number 34 (ENV/JM/MONO(2005)14). Paris, France.

Rispin, A., Harbell, J., Klausner, M. et al. (2004). Quality Assurance for In Vitro Test Methods: Quality Control Issues in Test Kit Production. In Proceedings of the Fourth World Congress on Alternatives and Animal Use in the Life Sciences. *ATLA 32, Supplement B*, 725-729.

Rispin, A., Stitzel, K., Harbell. J. and Klausner, M. (2005). Ensuring Quality of in vitro alternative test methods: current practice. *Reg. Tox. Pharm.* (in press).

U.S. Environmental Protection Agency (USEPA)(2002). 40 CFR 160 Good Laboratory Practice Standards (Office of Pesticides) and 40 CFR 792 Good Laboratory Practice Standards (Office of Prevention and Toxic Substances). Washington, D.C.

U.S. Environmental Protection Agency (USEPA)(2003). Paper entitled: Background for FIFRA Scientific Advisory Panel on Regulatory Use of In Vitro Systems. presented at FIFRA Scientific Advisory Panel Meeting, October 28-29, 2003. Arlington, Virginia.

U.S. Environmental Protection Agency (USEPA)(2004). Minutes of the FIFRA Scientific Advisory Committee Meeting held October 28-29, 2003. "A Set of Scientific Issues Being Considered by the Environmental Protection Agency Regarding: Ensuring Data Quality for In Vitro Tests Used as Alternatives to Animal Studies for Regulatory Purposes: A Consultation". SAP Minutes No. 2003-03, January 23, 2004. Arlington, Virginia.

U.S. Food and Drug Administration (USFDA)(2003). Center for Drug Evaluation and Research. 21 CFR 58 Good Laboratory Practice for Nonclinical Laboratory Studies. Washington, D.C.

Correspondence to
Amy S. Rispin, Ph.D.
U.S. Environmental Protection Agency
Office of Pesticide Programs (7506C)
Washington, D.C., USA
e-mail: rispin.amy@epa.gov

Session 5.08
Challenges in food toxicity testing

Safety Assessment of Genetically Modified (GM) Foods

Patrick O'Mahony and Alan Reilly
Food Safety Authority of Ireland, Dublin, Ireland

Summary
Modern biotechnology enables a transfer of genes between species that would not occur naturally. Safety assessment of genetically modified (GM) foods is primarily based on the concept of substantial equivalence, a comparison with conventional counterparts and similar varieties as a starting point. Nutritional and toxicological studies should be based on the outcome of the comparison, and additional safety tests should be carried out as required. The limitations of animal studies, particularly in assessing the safety of whole GM foods may be addressed by the new technologies that brought us GM foods in the first place. While animal studies may still have a role in the testing of GM food components, the advent of technology such as "genomics" offers a real and possibly superior alternative.

Keywords: modern biotechnology, genetically modified foods, substantial equivalence, genomics

Introduction

A genetically modified organism (GMO) is any living organism that has been manipulated at the level of its genome in a way that would not happen in nature. In theory, the DNA of any living organism can be manipulated, though most experience has been obtained through the genetic engineering of microorganisms (bacteria and viruses) and plants. However, a number of GM fish and GM animals have been developed and are awaiting authorisation for commercialisation. While microorganisms have been genetically manipulated to produce medicines, pharmaceuticals and food ingredients for some time, consumer opposition to GMOs became more widespread when crop plants were engineered to improve aspects of agricultural food production. While the reasons for this bias are multi-factorial, the food industry and regulators alike acknowledge the negative public perception and are endeavouring to improve the quality and transparency of food safety assessments. Such is the sensitivity to public perception that commercialisation of many modern biotechnology innovations in some parts of the world, including GM salmon, GM wheat and terminator technology, have been put on hold due to negative public opinion.

Science is an ever-evolving discipline and while the methods and means to produce new foods are being constantly renewed, so too must the tools used to assess the safety of new foods. The science that gave us modern biotechnology and GM food has also yielded technology that allows us to examine the safety of food at the molecular level. This new technology provides an option to reduce or possibly replace many of the animal studies routinely used in the safety assessment of GM foods where only minor differences from the parent plant are identified. An example of such a new methodology is toxigenomics, the combination of toxicology and genomics that allows the mapping of gene expression using microchip technology. Regardless of the methodology used to assess the safety of a food, a "completely safe" evaluation is not possible as no food is risk-free. This view is particularly salient given what we now know about the toxins, allergens and even carcinogens naturally present in many staple foods. In line with this, GM foods that are allowed on the EU market are classed

"as safe as their non-GM counterpart" without the guarantee of absolute safety, which does not even apply to the parent or non-GM variety.

Safety assessment of GM food

Guidelines for the safety assessment of GM food have been developed by international organisations (OECD, 1996; FAO/WHO, 1996, 2000, 2001, and 2002; and WHO, 2005), by the European Scientific Committee for Food (EU, 1997) and by the European Food Safety Authority (EFSA, 2004a). Standard GM food safety assessments are designed to identify characteristics such as nutrition, toxicity and allergenicity that can all impact on human health. Animal studies have played a significant role in GM food safety assessments and have been considered central to the identification of risk to human health. Animal studies are effective in identifying and quantifying the potential effects of chemical or microbial entities on a biological system where a cause and effect correlation is feasible. However, animal testing of whole foods such as maize or soya yields data of limited value, as the effects observed, if any, may be attributable to any of a number of the food constituents, while the consequences of minor constituents may go undetected. In addition, animal feeding studies are usually carried out on small animals such as rats, mice or chickens that have a relatively short life cycle, and the physical limitations of these animals may preclude testing for dose responses due to the potential for nutritional imbalance. Even using animals with a relatively short life cycle still requires weeks or months of tests, while biological and environmental variations demand that significant numbers of animals be tested with any results being subjected to statistical analysis. Assessment of the results of these tests can also lead to conflicting interpretations with respect to the significance of identified effects, as recently experienced by the European Food Safety Authority where the interpretation of a rat feeding study on a GM crop was challenged (EFSA, 2004b).

The majority of GM foods available on the market are derived from GM plants that differ from the parental variety only by the presence of one or a small number of introduced genes that ultimately result in the production of one or a small number of proteins or associated metabolites. Some examples are herbicide-resistant and/or insect-resistant soybeans, maize, cotton and oilseeds. The potential for introduced proteins possessing toxic or allergenic properties can be assessed *in vitro* and by using bioinformatics to compare the amino acid sequences of the new protein to known allergenic or toxic amino acid sequences stored in established databases. The insertion of a new gene into a plant genome can also result in unintended effects such as disruption in the expression or regulation of a plant gene. However, these unintended effects can generally be predicted and characterised through bioinformatic investigation of the insertion site. Unintended effects that are less predictable are possible however, and include interactions between an introduced protein and other proteins, DNA, RNA or metabolites, all of which may go unnoticed regardless of the method used.

Substantial equivalence

Fundamental to the international consensus on the safety assessment of GM food is the concept of substantial equivalence. First proposed by the OECD in 1993, (OECD, 1993) it has been developed and accepted as a first step in the safety evaluation of GM foods (FAO/WHO, 2004). The application of the principle of substantial equivalence is not a safety assessment in itself but is a framework for identifying similarities and differences between existing foods and new products. The properties of GM foods are compared with similar existing foods that have a long history of safe use, taking into account both intended and unintended effects. Detected differences are subjected to additional analyses to understand their significance and potential impact, which can ultimately be used in the final safety assessment. The OECD Task Force for the Safety of Novel Foods and Feeds, which first convened in 1999, develops and publishes "Consensus Documents" relating to important crop plants such as soybean, maize, potatoes, sugar beet and others (OECD, 2001-2005). These science-based consensus documents, which are mutually acceptable among Member States, contain what is considered to be key information on nutrients, anti-nutrients and toxicants for use in the regulatory assessment of new foods.

The concept of substantial equivalence has been challenged by various groups as vague and ill defined, and the term "substantial" has been misinterpreted in the past (Kok and Kuiper, 2003). These authors suggest that the principle should be rephrased as a "Comparative Safety Assessment", which describes the basic principle of comparing the new product with conventional counterparts, the relative safety of which are accepted.

Alternatives to animal testing for GM food safety assessment

Modern biotechnology has not only introduced new foods but also new techniques with which to characterise foods at the molecular level, which in turn provides an opportunity to reduce the reliance on animal testing. In particular, the advent of technology such as genomics, transcriptomics, proteomics and metabolomics offers new ways of examining the molecular details of food and thereby identifying even minor differences between GM plants and the non-GM parent. Genomics technology involves gene chips or microarrays that allow the analysis of thousands of genes at one time and which can identify genes that are expressed or switched off and whether expression is up or down regulated. Proteomics, including 2–dimensional polyacrylamide gel electrophoresis offers a direct window into a cell, revealing the presence or absence and levels of thousands of proteins on a single gel. Metabolomics utilises a number of standard techniques, such as GC, HPLC, MS and others, to analyse for and compare the levels of chemicals and metabolites. These techniques, when used to elucidate nutritional profiles, are termed "nutrigenomics" and for toxicology studies "toxigenomics", etc. This

technology presents advantages over animal studies in that a relatively short time scale is involved, it is more sensitive, being inclusive of minor constituents, and is reliable in that results are less reliant on statistical interpretation compared to animal studies. While genomics provides a more efficient comparative tool, it has also enabled the furnishing of a wealth of information to a number of accessible databases that can be interrogated to predict the toxicological or allergenic properties of uncharacterised proteins.

Evaluation of any significant differences and safety assessment

Once significant differences are identified between a GM plant and its non-GM parent, more detailed analyses of these constituent proteins or metabolites can be undertaken with a view of elucidating nutritional, toxicological or allergenic characteristics. While many of these studies can be carried out *in vitro* or using predictive bioinformatics, animal studies may be appropriate to determine dose responsiveness and empirical biological consequences. The significance of the observed differences can then be evaluated from the information provided at this stage and fed into the safety assessment, which may also take into account other factors such as potential intake.

Future GM foods

While the development and production of new GM foods within the EU is relatively insignificant, GM crop development and production in the rest of the world is progressing. Comparative safety assessment methodology, or substantial equivalence, could have a role to play in the safety assessment of GM foods into the future, even those with multiple added traits. However, as new, more complex and less benign GM foods are developed, assessing their safety may require a new approach. The engineering of food crops to produce additional nutrients, pharmaceuticals, medicines and even non-food constituents such as "bioplastics" promises to challenge the industry and regulators yet again to provide sufficient assurances about their safety to consumers.

Conclusions

The European Regulation (EC, 1996) on harmonised rules for the protection of animals used for experimental purposes obliges Member States to refrain from using animals where other testing methods are as effective and to promote the development of alternative experimental techniques. Numerous analytical techniques bases on "-omic" technologies exist or are under development and provide a new approach to understanding the effects of gene manipulation or chemical exposure. These advances will have significant implications for the safety assessment of GM foods and regulatory decision-making. Data generated from studies using toxigenomics, nutrigenomics and others will in future provide an objective way of assessing surrogate systems for reporting or predicting adverse effects of chemicals in humans that can replace or reduce the reliance on animal testing.

References

EC (1996). Council Directive 86/609/EEC of 24 November 1986 on the approximation of laws, regulations and administrative provisions of the Member States regarding the protection of animals used for experimental and other scientific purposes. *Official Journal L 358*, 18/12/1986. 0001-0028.

EFSA (2004a). Guidance Document of the scientific panel on genetically modified organisms for risk assessment of genetically modified plants and derived food and feed. *EFSA Journal (2004) 99*, 1-94.

EFSA (2004b). Opinion of the Scientific Panel on Genetically Modified Organisms on a request from the Commission related to the safety of foods and food ingredients derived from insect-protected genetically modified maize MON 863 and MON 863 x MON 810, for which a request for placing on the market was submitted under Article 4 of the Novel Food Regulation (EC) No 258/97 by Monsanto. (Question No EFSA-Q-2003-121).

EU (1997). Commission Recommendations 29 July 1997 concerning the scientific aspects and the presentation of information necessary to support applications for the placing on the market of novel foods and novel food ingredients and the preparation of initial assessment reports under Regulation EC 259/97 of the European Parliament and of the Council. *Off. J. Eur. Commun. L253*, 1–36. www.europa.eu.int/eur-lex/en/lif/dat/1997/en_397X0618.html.

FAO/WHO (1996). Biotechnology and Food Safety; Report of a Joint FAO/WHO Consultation, Rome, Italy, 1996. FAO Food and Nutrition Paper 61; Food and Agriculture Organisation of the United Nations, Rome.

FAO/WHO (2000). Safety Aspects of Genetically Modified Foods of Plant Origin. Report of a Joint FAO/WHO Expert Consultation on Foods Derived from Biotechnology, Geneva, Switzerland, 2000, Food and Agriculture Organisation of the United Nations, Rome, www.fao.org/es/esn/food/gmreport.pdf.

FAO/WHO (2001). Allergenicity of Genetically Modified Foods. Report of a Joint FAO/WHO Expert Consultation on Foods Derived from Biotechnology. Rome, 2001, Food and Agriculture Organisation of the United Nations, Rome, www.who.int/fsf/Documents/Biotech_Consult_Jan2001/report20.pdf.

FAO/WHO (2002). Report of the Third Session of the Codex Ad Hoc Intergovernmental Task Force on Foods Derived from Biotechnology (ALINORM 01/34). Codex Ad Hoc Intergovernmental Task Force on Foods Derived from Biotechnology, Food and Agriculture Organisation of the United Nations, Rome, www.fao.org/codex/alinorm03/Al03_34e.pdf).

FAO/WHO, (2004). *Codex Alimentarius – Foods derived from modern biotechnology*. Rome: Food and Agriculture Organisation of the United Nations, ISBN 92-5-105259-X.

Kok, E. J. and Kuiper, H. A. (2003). Comparative Safety Assessment for biotech crops TRENDS. *Biotechnology 21(10)*, 439-444.

OECD (1993). Safety Evaluation of Foods Derived by Modern Biotechnology: Concepts and Principles. Paris: Organisation for Economic Co-operation and Development. www.oecd.org/pdf/M00034000/M00034525.pdf.

OECD (1996). Food Safety Evaluation. Paris: Organization for Economic Cooperation and Development.

OECD (2001-2005). Consensus Documents for the work on the Safety of Novel Foods and Feeds. Paris: Organization for Economic Cooperation and Development. http://www.oecd. org/document/9/0,2340,en_2649_201185_1812041_1_1_1_1,00.html.

WHO (2005). Modern Food biotechnology and development: an evidence-based study. Geneva, Switzerland: World Health Organisation. www.who.int/foodsafety/biotech/who_study/en/index.html.

Correspondence to

Patrick O'Mahony
Food Safety Authority of Ireland
Abbey Court, Lower Abbey Street
sDublin 1, Ireland
e-mail: pjomahony@fsai.ie
http://www.fsai.ie

Session 5.09
Biologicals and vaccines: Progress and new approaches

Development of Alternatives for Quality Control of Biomedicines: The Example of Vaccines for Human Use

Marie-Emmanuelle Behr-Gross
European Directorate for the Quality of Medicines, Council of Europe, Strasbourg, France

Summary
The European Directorate for the Quality of Medicines (EDQM) co-ordinates the elaboration of the European Pharmacopoeia. Within the field of biomedicines, the EDQM has been able to implement high standards due to the contribution of experts and to the European Biological Standardisation Programme (BSP). The aim of the BSP is to establish reference methods and preparations with a focus on alternatives to animal testing: potency assay validation studies for vaccines illustrate the potential contributions of the BSP to animal welfare. The successful projects run for diphtheria and tetanus vaccines will be extended. Based on the assumption that sera from the same animals (guinea pigs) could be used for in vitro determination of antibodies to different components of combined vaccines, the potential for reduction and refinement of the routine vaccines batch potency assay by this approach is substantial and will be further explored by the BSP.

Keywords: European Pharmacopoeia, quality control, biomedicines, vaccines, alternatives

Introduction

Marketing of medicines is regulated in the European Union (EU) by the relevant directives[1]. Medicines quality standards in 35 countries and organisations[2] – including the EU –, which are parties to the European Pharmacopoeia (*Ph. Eur.*) Convention[3], are defined in monographs and general methods elaborated by medicines control experts groups under the co-ordination of the EDQM. New or revised monographs and general methods are submitted to public enquiry through publication in the European pharmacopoeial forum "Pharmeuropa" and then endorsed by the *Ph. Eur.* Commission (representatives of parties to the Convention) prior to being published and implemented. Adherence to the "European Convention for Protection of Vertebrate Animals used for Experimental and Other Purposes"[4] in the field of medicines Quality Control (QC) is strongly promoted by the EDQM[5].

Among the products covered by the *Ph. Eur.* are many biologicals which require extensive QC testing to demonstrate their quality, safety and efficacy. High quality standards for biomedicines have been implemented in Europe thanks to contributions of specialised *Ph. Eur.* experts groups and to the development of the Biological Standardisation Programme

[1] European Commission directives : 2004/27/EC amending 2001/83/EC, 2004/24/EC amending 2001/83/EC and 2004/28/EC amending 2001/82/EC
[2] see www.pheur.org
[3] Council of Europe European Treaty Series No. 50, 22/07/1964 and No. 134. 01/11/1992
[4] Council of Europe European Treaty Series No. 123, 18/03/1986
[5] Artiges, A. (1999). Alternatives to animals in the development and control of biological products for human and veterinary use. The role of the European Pharmacopoeia. *Dev. Biol. Stand. 101*, 29-35.

(BSP), a research programme co-sponsored by the EU (European Commission Directorate General-Enterprise, Pharmacy Division) and the Council of Europe.

The aim of the BSP is to establish reference methods and materials for biomedicinal products with a special emphasis on validation of alternatives to animal testing. Division IV of the EDQM is responsible for co-ordination of the BSP.

This manuscript focuses on potential contributions of the BSP to animal welfare issues in the field of vaccines for human use regarding serological potency assay validation (table 1). Two projects, coded BSP019 and BSP035, were run using a guinea pig serological potency assay model for tetanus (T) vaccines for human use. A follow-up project (BSP034) dealt with diphtheria (D) vaccines and provided confirmation of the assumption that sera from the same animals could be used for *in vitro* determination of antibodies to both T and D components of combined vaccines. Future plans include the validation of ELISA methods for an acellular pertussis vaccine serological potency assay in guinea pigs.

Features of the T vaccine potency assay validation studies will be described here-in. Indeed these projects were pilot studies from which it was possible to identify and resolve a number of issues. These projects also demonstrated an exemplary collaboration with the EU through the sponsorship provided by the European Centre for Validation of Alternative Methods (ECVAM-IHCP-JRC) to pilot laboratories, a good model for effective regulatory acceptance and implementation of alternatives through *Ph. Eur.* general methods revisions and a test as regards the potential for harmonisation of international medicines QC standards in line with the European animal welfare policy.

The tetanus studies were divided into 2 projects, BSP019 and BSP035, which were initiated in 1996 and completed in 2000. These projects were conducted under the scientific direction of Dr. C. Hendriksen (Netherlands Vaccine Institute, NL) and Dr. R. Winsnes (Norwegian Medicines Agency, N) and co-ordinated by EDQM.

A BSPO19

A.1 Context

In 1996, the tetanus potency assay prescribed by the *Ph. Eur.* 4th edition was a quantitative direct challenge test in guinea pigs or mice (immunisation followed by toxin challenge), generating severe distress and pain in animals. The replacement of this test was therefore considered a priority for the BSP, which was supported by ECVAM-IHCP-JRC in this endeavour. Potential alternatives included serology, where the toxin challenge is replaced by an antibody detection test, or pure *in vitro* tests, e.g. antigen content measurement of the vaccine.

A.2 Candidate alternative method selection

In order to determine the method of choice for validation as an alternative to challenge, the following criteria were applied.

A Candidate method should:

- represent an animal welfare and technical improvement.
- be validated for use in QC in at least one laboratory, transferable to other laboratories.
- be based on Standard Operating Procedures (SOPs) and reagents available to EDQM and to potential users in the future.
- make use of established reference preparations, if relevant.
- have the potential to be accepted further to validation by regulatory authorities in view of implementation as an official standard in Europe and elsewhere to promote international harmonisation.

Analysis according to the criteria listed in A.1 revealed that only the serological potency assay complied with these prerequisites.

A.3 Animal model and *in vitro* method selection

In order to determine a suitable animal model and an *in vitro* method for antibody measurement, the following criteria were applied.

An animal model should:

- mimic the situation in humans.
- demonstrate the relationship between challenge and antibody response with documented variability sources (e.g. strain, breeding conditions, sex, etc.).
- generate sufficient amounts of sera to assay several components of combined vaccines on the same animals.

An *in vitro* method should:

- be validated (in-house).
- be easily transferable to other QC laboratories.
- be evaluated for vaccine QC, notably for its ability to discriminate between good and borderline vaccines.

The Toxin Binding Inhibition Test (ToBI) and Enzyme Linked Immunoassay (ELISA) for determination of tetanus antibodies in guinea pigs complied with the prerequisites set by organisers and project leaders.

Tab. 1: BSP projects for human vaccines serological potency assay validation

Product	Method	Monograph(s) – General Method	Project(s)
Acellular pertussis vaccine	Assay (serological method(s) based on ELISA	2.7.16. Assay of pertussis vaccine (acellular)	BSP083
Diphtheria vaccine (adsorbed)	Assay (serological method(s) based on Vero cell test and/or ELISA	2.7.6. Assay of diphtheria vaccine (adsorbed)	BSP006 BSP034
Tetanus vaccine (adsorbed)	Assay (serological method(s) based on ELISA and ToBI)	2.7.8. Assay of tetanus vaccine (adsorbed)	BSP019 BSP035

A.4 Organisation of the project – objectives

Further to the selection of potential alternatives to the challenge assay, the study was divided into different phases, aimed at addressing successive goals.

A.4.1 Prevalidation phase

The aim of the prevalidation phase was to optimise the methods and test design to suit the purpose of BSP019. Optimal immunisation doses and period, bleeding methods, ELISA and ToBI procedures were defined in one laboratory.

A.4.2 Phase I

All methods were transferred to a second laboratory, which confirmed the results obtained formerly, using the same testing procedures and design.

A.4.3 Phase II

In Phase II, 5 laboratories were requested to test a set of vaccines including 1 borderline vaccine in challenge and by using the candidate alternatives. Correlation between potencies (*in vivo* vs. serology) and between antibody titres and protection in the individual animal were shown.

A.5 Issues and knowledge acquired from Phases I and II

A number of issues were raised and solved during phases I and II. Amongst them, it is useful to take note of the following observations.

A.5.1 Number and experience of testing laboratories

For obvious animal welfare reasons, it is recommended to limit the numbers of animals used for *in vivo* testing in alternatives validation projects. It is however recommended to enrol a sufficient number of very experienced laboratories to avoid delays and repetition of animal experiments due to test failure (e.g. BSP019, phase IIa, 3 labs: failure of *in vivo* test in 1 laboratory and of *in vitro* test in another laboratory, therefore extension to 2 more laboratories needed, phase IIb).

A.5.2 Borderline products and selection of test vaccines.

The testing of borderline products is usually not possible, as such products are frequently unavailable. It is generally accepted that artificially weakened products (e.g. by dilution) are used instead. In BSP019 however, a borderline vaccine could be studied, in parallel with 6 other vaccines complying with *Ph. Eur.* potency specifications.

A.5.3 Test-scheme

To generate appropriate dose-response relationship in the serological model, it is necessary to explore the behaviour of the method using a multi-dilution immunisation scheme, perhaps with more doses than would be used in routine QC. For *in vivo* testing, where difference in challenge toxin, strains and origin of animals and conditions of breeding may interfere, test-schemes must be elaborated on an individual basis for each laboratory.

A.5.4 Sourcing of essential reagents

Sourcing of non-commercial reagents needs to be addressed ahead of the start of a collaborative study. Sufficient amounts must be produced and made available to the EDQM.

A.5.5 Statistical analysis

The impact of the use of different software and statistical models on estimated parameters must be evaluated, despite the labour-intensive approach this implies.

In BSP019 the following models were used:

- probit analysis for challenge data.
- 4 and 5 parameters models, dichotomisation and probit analysis for tetanus antitoxin concentrations.
- logistic regression for protective concentration.
- probit analysis after dichotomisation.
- Sign test and correlation coefficients (Pearson) were used for correlation.

A.5.6 Continuation of study

Critical evaluation of the first phases enabled making the decision on whether to extend to the large international validation study.

B BSP035

B.1 Context

In view of the duration and large amount of data generated in phases I and II of BSP019, it was decided to start a new project to run an enlarged international collaborative study for the validation of the *in vitro* part of the serological potency assay. This part of the study is referred to as Phase III.

B.2 Aim

The aim of Phase III was to assess intra- and inter-laboratory variations in estimation of T-antibodies by ELISA and ToBI.

B.3 Issues and knowledge acquired from Phase III

B.3.1 Participants

Enrolment of participants experienced with the test methods is deemed necessary. This implies that both public Official Medicines Control Laboratories (OMCLs) and private (manufacturers) sector QC laboratories participate on a geographical basis that may be larger than Europe. In BSP035, 25 laboratories from 13 European countries, USA, Australia and India took part.

B.3.2 Test samples

Production of a large panel of antisera implies knowledge of marketed products and support from manufacturers for sourcing of test vaccines and antisera. A choice has to be made among available vaccines. In BSP035, 14 vaccines and references, corresponding to different components and potencies, enabled production of a panel of 28 sera.

B.3.3 Essential reagents

Identification of suppliers of essential reagents for the collaborative study and for future validation studies at QC laboratories is needed. For BSP035, RIVM (NL) and NIBSC (UK) supplied EDQM with essential reagents for ToBI and ELISA. A stock of essential reagents for ToBI was made available to the EDQM by RIVM further to completion of the study and a tetanus toxoid batch suitable for ELISA and ToBI was donated by SSI (DK).

B.3.4 Standardisation of *in vitro* tests

It is useful to determine the extent of standardisation needed in the study and also for subsequent use in routine QC: in general this can be addressed by prescribing use of in-house procedures in parallel to centrally provided SOPs. In BSP035, only a few laboratories performed in-house tests. After completion of BSP035, SOPs and essential reagents were therefore provided to all potential users upon request.

B.3.5 Analysis of study data

In most collaborative studies, calculations are to be performed by participants and repeated in a central statistical analysis.

However, in case of validation of a new method, this may not be required. In BSP035, only a central analysis was performed. Based on the conclusion of this analysis, a suitable statistical model was prescribed for use of the test in routine QC.

B.4 Conclusions of Phase III

Several conclusions could be reached in BSP035:

- In terms of test reliability: ELISA vs. ToBI generate different results that may indicate that there are differences in the antibodies measured.
- In terms of test suitability: ELISA and ToBI are both considered suitable for monitoring tetanus antitoxin levels in guinea pig sera.
- In terms of suitability for vaccines batch release: ELISA and ToBI-based serological methods are suitable, provided that the test is carefully designed and variability of different steps is monitored.

C Implementation of serological potency assay as a regulatory test for tetanus vaccine for human use

The adoption of the newly validated potency assay as a *Ph. Eur.* standard was the result of several steps subsequent to the collaborative studies.

C.1 Communication

A symposium took place on 22-23 June 2000 to address the need for changes in T vaccine QC standards[6]. Publication of the reports of the collaborative studies in the Pharmacopoeial forum[7] allowed the results of BSP019 and BSP035 to be shared widely with the scientific community and the regulatory authorities worldwide. The project leaders and EDQM experts and representatives also participated in scientific and WHO meetings on vaccines standardisation.

C.2 Technical support to potential users

SOPs and essential reagents for ToBI and ELISA are provided upon request to manufacturers and OMCLs.

C.3 Inclusion in *Ph. Eur.* Group 15 work programme

After completion of BSP035, the group of 15 experts of the *Ph. Eur.* started to revise the general method 2.7.8. Assay of tetanus vaccine (adsorbed), resulting in the inclusion in 2001 of a new method referred to as "method C: measurement of antibodies in guinea pigs". The revised text was published for public enquiry in Pharmeuropa 14.3 in July 2002, and adopted by the *Ph. Eur.* Commission in March 2003. The new method was published in the *Ph. Eur.* (4.7) in September 2003.

D Promotion of reduction and refinement of animal testing for tetanus vaccines, follow-up studies

Humane end-points and single dilution design of challenge constituted a step forward towards the 3Rs. Moving from challenge to serological potency assay represents another significant step, but to encourage manufacturers and OMCLs to move to this test, economical and technical incentives must be found. The BSP was therefore asked to look into the possibility of testing several components on the same sera. Most vaccines currently marketed are combined, and this is currently considered the most promising approach. Follow-up studies have been undertaken since 2001 and are listed below.

D.1 BSP034

BSP034, a project on D vaccine serological potency assay started in 2001 and was completed in 2004. A Vero cell assay and an ELISA-based serological potency assay were validated. By demonstrating that T and D antibody contents can be measured in sera from the same guinea pigs, the study has shown that determining potency for D and T components using a refined test requiring fewer animals was possible[8,9].

Currently, method 2.7.6. Assay of diphtheria vaccine (adsorbed), is undergoing revision. The public enquiry started with publication *in Pharmeuropa* 17.3 in July 2005.

D.2 BSP083

BSP083 started in 2005 with the aim of validating ELISAs for detection of antibodies to acellular pertussis components (pertussis toxin, fimbrial haemagglutinin, pertactin, Fim. 2/3 antigens).

D.3 Perspectives

Scientific data still needs to be generated in order to determine whether the guinea pig "same animal" model can also be used to monitor batch potency for other components of combined vaccines, e.g. whole cell pertussis, poliomyelitis and hepatitis B vaccines.

E General conclusion

Performance of collaborative validation studies BSP019, BSP034 and BSP035, involving manufacturers and authorities on a worldwide basis, was initiated by the EDQM in the field of human vaccines, with obvious positive consequences for implementation of 3R-compliant QC policies. Experience from the BSP benefits new projects and has enabled the introduction of alternatives in the *Ph. Eur.* BSP projects have also contributed to international harmonisation, by allowing non-European authorities and the WHO to be proactively involved at all stages of the projects.

F Acknowledgement to BSP019, BSP034 and BSP035 contributors

Study participants are warmly thanked for their ongoing participation. Dr. C. Hendriksen, Dr. R. Winsnes, and Dr. D. Sesardic were the project leaders. Dr. A. Akkermanns and A. Daas performed statistical analysis. Special thanks go to Dr. D. Sesardic (NIBSC, UK), Dr. K. Haslov (SSI, DK) Dr. J. W. Petersen (SSI, DK), Dr. J. Van der Gun (NVI, NL), Dr. M. Duchêne (GSK, B), Dr. O. Van Loocke (GSK, B), Dr. A. Sabouraud (Sanofi-Pasteur,

[6] Pharmeuropa (Oct 2000). International Symposium, Strasbourg, 22-23 June 2000. Tetanus vaccines for human use. Proceedings of the Symposium

[7] Pharmeuropa Bio (Oct. 2001). Special Issue Biol. Bio 2001-2.

[8,9] Sesardic, D., Winsnes, R., Rigsby, P. and Behr-Gross, M-E. (2003). Collaborative study for the validation of serological methods for potency testing of diphtheria vaccines – extended study: correlation of serology with *in vivo* toxin neutralisation. *Pharmeuropa Bio. 2004*, 69-76.

Winsnes, R., Sesardic, D., Daas A. and Behr-Gross, M-E. (2003). Collaborative Study for the Validation of Serological Methods for Potency Testing of Diphtheria Toxoid Vaccines – Part 1. *Pharmeuropa Bio. 2004,* 35-68.

F), Dr. N. Hug (Chiron Vaccines, I) and Dr. St. Knapp (Chiron Behring GmbH, D) for donation of reagents and test samples.

G Concluding remarks

Experience gained in the BSP programme contributes to revising and updating the regulatory frame of biomedicines in Europe in line with the European Convention for the Protection of Vertebrate Animals for Experimental and other Scientific Purposes[4]. However, additional efforts are needed, notably at the level of policy makers, to promote research and development on alternative methods and international harmonisation for medicines control in Europe.

Nevertheless, a number of 3R issues linked to QC control of biomedicines are out of the BSP scope. Indeed for many products:

- harmonisation of regulations in different regions has not yet been achieved.
- implementation of *in vitro* methods is dependant on economical parameters (e.g. in the developing countries animal tests are cheaper than *in vitro* tests).
- lack of resources to move to new tests is frequently reported, notably by public sector laboratories.
- the start of new studies is strongly hindered by a lack of knowledge and research into potential alternatives.

To promote the 3Rs in medicines control, the cooperation of policy makers, public health and regulatory authorities, standardisation bodies, medicines control experts, scientists as well as animal welfare organisations is sought. In particular, there is an absolute need for research and development and potential sponsors have to be identified amongst OMCLs, manufacturers, national and international research platforms, European Commission (FP6 Thematic priority 1), ECVAM-IHCP-JRC and the various animal protection organisations.

Correspondence to

Dr. Marie-Emmanuelle Behr-Gross
European Directorate for the Quality of Medicines
Council of Europe
226, Avenue de Colmar
67029 Strasbourg
France
e-mail: marie-emmanuelle.behr-gross@pheur.org

Contributions of the European OMCL Network and Biological Standardisation Programme to Animal Welfare

Marie-Emmanuelle Behr-Gross and Jean-Marc Spieser
European Directorate for the Quality of Medicines, Council of Europe, Strasbourg, France

Summary
In 1991 a contractual co-operation in harmonising medicines control started between the European Union and the European Directorate for the Quality of Medicines (EDQM)-Council of Europe (CoE): The EDQM has been charged to co-ordinate a network of national Official Medicines Control Laboratories (OMCL) and a research programme referred to as Biological Standardisation Programme (BSP). In line with the CoE convention on the protection of animals, the BSP establishes European Pharmacopoeia (Ph. Eur.) standards for biomedicines quality control with a special emphasis on 3Rs alternative methods. Sixteen projects on vaccines and one on blood products have been initiated in this field. The programme, run in the spirit of international harmonisation, involves the OMCL network, public and private sector medicines control laboratories in Europe, the Americas, Asia and Australia and non-European standardisation bodies. Completed projects on Newcastle disease and clostridial veterinary vaccines and on diphtheria and tetanus human vaccines led to new Ph. Eur. general methods and standards, thus showing that the BSP promotes regulatory acceptance of alternatives. Further studies deal with botulinum toxin, vaccines for human use (inactivated poliomyelitis virus, hepatitis A, hepatitis B and pertussis) and tetanus immunoglobulin. For the future, the programme hopes to benefit from synergies between fundamental, medical and pharmaceutical sciences experts for promoting animal welfare aspects in control whilst guaranteeing quality, safety and efficacy to biomedicines potential users.

Keywords: biological standardisation programme, alternative method development, three Rs, validation acceptance, biomedicinal products, routine release, quality control

The BSP programme and its achievements

A contractual co-operation between the European Union and the European Directorate for the Quality of Medicines (EDQM)-Council of Europe (CoE) started in November 1991 with the aim of improving harmonisation within the European framework for the control of medicines. For that purpose, the EDQM has been charged with setting up and co-ordinating the activities of a network of national Official Medicines Control Laboratories (OMCL) (Paulsen-Sörmann et al., 2000; Milne et al., 2002) and to develop a scientific research programme referred to as the Biological Standardisation Programme (BSP). In agreement with the Council of Europe convention on the protection of animals[1] and with the general policy adopted by the European Pharmacopoeia (*Ph. Eur.*) (Artiges, 1999), the OMCL network has committed itself to promoting the 3Rs (Russell and Burch, 1959), in particular in the field of biomedicines control. The BSP, supported by scientific advisors from this network, establishes *Ph. Eur.* standard preparations and methods with a special emphasis on validation of "alternative" methods for reduction, refinement and replacement of animal testing. Public and private sector medicines control laboratories from Europe, the Americas, Asia and Australia participate actively in the BSP projects thus enabling

a) to successfully validate alternative methods for biomedicines quality control purposes
b) to enhance regulatory acceptance of the validated alternative methods
c) to establish reference materials (reagents or Biological Reference Preparations) for alternative methods specified in the *Ph. Eur.*
d) to promote an internationally harmonised scientific and ethical approach to biomedicines control.

17 projects dealing with "3Rs-compliant" control methods or standards for human vaccines (13), veterinary vaccines (3) and blood products (1), have been initiated within this framework. Integration of new general methods and adoption of new standards by the European Pharmacopoeia have taken place after successful completion of projects; amongst those the following are the most recent achievements: new methods for the control of Newcastle disease (Claasen et al., 2004) and clostridial veterinary vaccines (Lensing et al., 2000; Lucken et al., 2002) and for diphtheria (Sesardic et al., 2003; Winsnes et al., 2003) and tetanus[2,3,4] vaccines for human use. Ongoing projects deal with various topics: botulinum toxin, human vaccines (inactivated poliomyelitis virus, hepatitis A, hepatitis B and pertussis) and tetanus immunoglobulin. Whenever possible, BSP contributions and plans for 3Rs are communicated to the life sciences community through publications and at meetings. Indeed, synergies between fundamental, medical and pharmaceutical sciences are

[1] European Convention for the Protection of Vertebrate Animals used for Experimental and other Scientific Purposes. Council of Europe, *ETS* 123, 18.III.1986.

deemed indispensable for promoting animal welfare aspects in biomedicines control whilst guaranteeing quality, safety and efficacy for potential users.

Experiences from the last 20 years of developing alternatives to animal testing by EDQM

It should be noted that the methods used for the control of medicines are subjected to very strict procedures defined by the Codes for human[5] and veterinary[6] medicines, including detailed validation, independent assessment, authorisation and supervision of their application. Any change in the authorised methods must follow these procedures, and demonstration of equivalency in terms of quality, safety and efficacy is a prerequisite. Consequently, it is of prime importance that all the actors concerned, i.e. manufacturers, control authorities and regulators, are involved at an early stage in a close collaboration to achieve an efficient and cost saving development of alternative methods. This development has to be based on experimental work with the support and active contribution of all concerned.

The international dimension as well as the inclusion in the studies of so-called "failing" batches (subpotent, impurity loaded, etc.) should always be considered when defining the project plan. Any other approach of the issue will require unnecessary duplication of validation and acceptance.

A clear separation should be made between routine Quality Control (QC) for batch release and investigational work aimed at demonstrating biological activity and safety aspects.

EDQM, through its function of European Standards setting body and its privileged and officialised liaison with the national and international agencies and institutions regulating medicinal products (EMEA, FDA, WHO, PEI, NIBSC, AFSSAPS, etc.), has a unique role to play in the setting up and implementation of alternative methods to animal testing.

The development of alternatives to animal testing is closely linked with the technical development in the domain of production of biologicals, which nowadays includes highly sophisticated purification and biotechnological processes. The application of detailed and controlled working procedures based on good manufacturing practices (GMP) throughout the entire production chain, from the starting material to the final product and its individual packaging, contributes also to a higher quality of biologicals.

Furthermore, technical developments in the field of analytical procedures should enable animal free or animal friendly QC tests based on molecular biology, immunological/immunochemistry and separative methods to be implemented.

This is now current practice at the R & D laboratory level, especially within the EU. However, implementation of such methods as routine practices through appropriate scaling up (e.g. 1 million individual doses of packaged vaccines) requires enormous investment and full validation at each stage of the industrial scaling up.

[2] Collaborative Study for the Validation of Serological Methods for Potency Testing of Tetanus Toxoid Vaccines for Human Use. Part 1. *Pharmeuropa Bio.* 2000 Aug;2000(1):83-124.

[3] Collaborative Study for the Validation of Serological Methods for Potency Testing of Tetanus Toxoid Vaccines for Human Use. Summary of all three Phases. *Pharmeuropa Bio.* 2001 Oct;2001(2):73-9.

[4] Collaborative Study for the Validation of Serological Methods for Potency Testing of Tetanus Toxoid Vaccines for Human Use. Part 2. *Pharmeuropa Bio.* 2001 Oct;2001(2):45-72.

[5] Code for human medicines: Directive 2001/83(EC) as amended by Directive 2004/27(EC) and 2003/63(EC)

[6] Code for veterinary medicines: Directive 2001/82(EC) as amended by Directive 2004/28(EC)

References

Artiges, A. (1999). Alternatives to animals in the development and control of biological products for human and veterinary use. The role of the European Pharmacopoeia. *Dev. Biol. Stand. 101*, 29-35.

Claassen, I., Maas, R., Oei, H. et al. (2004). Validation Study to Evaluate the Reproducibility of a Candidate In Vitro Potency Assay of Newcastle Disease Vaccines and to Establish the Suitability of a Candidate Biological Reference Preparation. *Pharmeuropa Bio. 2004 Dec. 2004(1)*, 1-16.

Lensing, H. H., Esposito-Farese, M.-E., Daas, A. and Spieser, J.-M. (2000). Collaborative Study for the Establishment of two European Pharmacopoeia Biological Reference Preparations for Clostridium Tetani Antiserum for Serological Potency Testing of Tetanus Vaccines for Veterinary Use. *Pharmeuropa BIO. 2001 Feb. 2000(2)*, 43-54.

Lucken, R., Daas, A. and Esposito-Farese, M.-E. (2002). Collaborative Study for the Establishment of a European Pharmacopoeia Biological Reference Preparation for Clostridia Antiserum for Serological Potency Testing of Clostridial Vaccines for Veterinary Use. *Dev. Biol. 111*, 171-180.

Milne, C., Buchheit K.-H. and Spieser, J.-M. (2002). The EU/EEA Network for Official Control Authority Batch Release of Biologicals for human use. *RAJ Pharma 13(6)*, 477-483.

Paulsen-Sörman, U., Wanko, R. and Spieser, J.-M. (2000). Market surveillance of Centrally Authorised Products, *RAJ Pharma 11(6)*, 409-413.

Russell, W. M. S. and Burch, R. L. (1959). *The principles of humane experimental technique*. London: Methuen & Co., Ltd.

Sesardic, D., Winsnes, R., Rigsby, P. and Behr-Gross, M.-E. (2003). Collaborative study for the validation of serological methods for potency testing of diphtheria toxoid vaccines – extended study: correlation of serology with in vivo toxin neutralisation. *Pharmeuropa Bio. 2004 Jan. 2003(2)*, 69-76.

Winsnes, R., Sesardic, D., Daas, A. and Behr-Gross, M.-E. (2003). Collaborative Study for the Validation of Serological Methods for Potency Testing of Diphtheria Toxoid Vaccines – Part 1. *Pharmeuropa Bio. 2004 Jan. 2003(2)*, 35-68.

Correspondence to

Jean-Marc Spieser
European Directorate for the Quality of Medicines - Council of Europe
226, Avenue de Colmar
67029 Strasbourg, France
e-mail: jean-marc.spieser@pheur.org

USDA 3Rs Initiatives in Veterinary Biologics

Jodie A. Kulpa-Eddy
United States Department of Agriculture (USDA)-Animal and Plant Health, Inspection Service (APHIS)-Animal Care, Riverdale, USA

Summary
This paper provides a brief overview of the U.S. system of regulatory oversight of the production of veterinary biologicals (vaccines, bacterins, diagnostic test kits intended for use in animals). Alternative test methods have been accepted by the U.S. Department of Agriculture (USDA), Animal and Plant Health Inspection Services (APHIS), Veterinary Services, Center for Veterinary Biologics (CVB). Both historical and current initiatives in this area are discussed.

Keywords: veterinary biologics, alternative test methods

Background

Federal regulation of veterinary biologics in the United States began in 1913 with passage of the Virus-Serum-Toxin Act. This law was enacted largely because of public concern over the importation of contaminated veterinary vaccines from Europe and in reaction to complaints about worthless and contaminated hog cholera products being sold throughout the country. This law requires the Department of Agriculture to ensure that veterinary biologics (vaccines, bacterins, antiserums and similar products) sold in interstate commerce are pure, safe, potent, and efficacious (CVB “Background”).

For nearly 50 years, the biologics programme was carried out by veterinary field inspectors located in the commercial biologics manufacturing establishments and by programmeme staff in Washington, DC. In 1961, the biologics programmeme was allocated ten percent of the space at the National Animal Disease Laboratory (now Center) that had just been established in Ames, Iowa. Since October 1996, the veterinary biologics programme has been carried out by the Center for Veterinary Biologics (CVB), a unit of Veterinary Services, APHIS, USDA. There are three units of the CVB: Laboratory; Licensing and Policy Development; and Inspection and Compliance (CVB "Background”).

“3Rs Initiative”

The Center for Veterinary Biologics recognises the responsibility to minimise animal pain and distress in research and testing procedures. In 2003, special emphasis was placed on the consideration of alternatives when the following goal was included in the strategic plan: to “significantly refine, replace and reduce animal testing of veterinary biologics” (CVB “Strategic Goals”). These “3Rs" must be accomplished to the degree possible while maintaining credible animal disease and immunity models to correlate *in vitro* and *in vivo* assays of the safety, potency and purity of veterinary biologics.

Replacement

In the 1960s and early 70s, all vaccine serial releases required the vaccination and challenge of target species or surrogate laboratory animals. This testing required a large number of animals to experience pain or distress in order to prove the vaccine released for use by the public was effective. That is, the control animals must succumb to the disease while the vaccinated animals must survive.

The “master seed” principle introduced in the 1970s was a major step towards reducing animal use and is the first example of *in vitro* potency testing. The vaccinate/control challenge potency test for modified live virus (MLV) vaccines was replaced by quantification of the live organisms (titration). Given that approximately half of the regulated biologics products are MLV vaccines, this resulted in an almost 50% decrease in animal use.

Today CVB is evaluating enzyme-linked immunosorbent assays (ELISAs) to determine their usefulness as replacement tests for animal vaccinate/control challenge potency tests. An ELISA test of specific interest is one that measures the presence of a specific protective antigen found on *Leptospira* bacteria. CVB has contracted with a researcher at Michigan State University to determine the viability of the ELISA system as a replacement for the more expensive, time-consuming and painful/distressful hamster potency test.

In addition, CVB is considering new regulations to allow in-process potency testing for serial release. This is being done to take advantage of *in vitro* tests that might not otherwise be possible due to adjuvant and other ingredient interference. These tests hold the potential for improved reproducibility and reliability. There will be some additional tests required to address qualitative and quantitative issues, such as verifying identity and formulation on the final product, but these too would be *in vitro* tests.

Reduction

A significant undertaking in the area of reduction is occurring under the auspices of the Office International des Epizooties

(OIE) through the International Cooperation on Harmonisation of Technical Requirements for Registration of Veterinary Medicinal Products (VICH). Officially launched in 1996, the third public conference was recently held in Washington, DC in May 2005. Members include the European Union, the U.S. and Japan; Canada, New Zealand and Australia participate as observers.

One of the objectives is to establish and implement harmonised regulatory requirements which meet high quality, safety and efficacy standards and minimise the use of test animals and costs of product development. Thirty-four guidelines have been finalised to-date. An example in the area of animal test requirements is the guideline developed for "Target Animal Safety" (VICH GL41) (FR 2004). It was published for public comment in the U.S. in December 2004. Comments received will be reviewed by the Expert Working Group, and then submitted to the VICH Steering Committee for finalisation and implementation. Once accepted, this will allow manufacturers to conduct one standard test that all the member countries have agreed is adequate for the given purpose.

Refinement

Two recent notices provide examples of acceptance of refinement techniques in biologics testing by CVB.

In April 2004, CVB issued notice of their policy concerning the use of humane endpoints in animal challenge tests. While the regulations still appear to indicate the animals must die from the challenge in order to be considered a valid test, CVB has modified their interpretation to state: "Moribund animals exhibiting clinical signs consistent with the expected disease pathogenesis that are unable to rise or move under their own power may be humanely euthanised and considered as deaths." Death is no longer a required endpoint (CVB Notice 04-09).

This notice also contains the acceptance of an early endpoint for the rabies vaccine challenge test: "Animals exhibiting paresis, paralysis, and/or convulsions may be humanely euthanised and considered as deaths."

More recently the CVB issued a notice regarding "Mouse Safety Testing". This is an announcement that CVB is considering revision of the test requirement outlined in Title 9 of the Code of Federal Regulations, Section 113.33(a)(1) for live virus vaccines. Currently the test requires inoculation of two groups of mice, one intracerebrally and the other intraperitoneally. CVB is proposing to revise the regulation to require that only one group of mice be inoculated, by either the intraperitoneal or subcutaneous route. The potential for pain or distress to the animal by these routes of injection is much lower. Until the regulation is revised, CVB will allow exemptions to the requirement on a case-by-case basis (CVB Notice 05-01).

Future Endeavours

In April 2004, CVB collaborated with the Institute for International Cooperation in Animal Biologics to sponsor a meeting on "Technology and Approaches to Reduce, Refine and Replace Animal Testing". The objective of this meeting was to discuss policy, provide guidance, and highlight new technologies with the potential for use in developing *in vitro* assays for vaccine potency, consistency, and stability. New approaches to measuring the quantity and quality of antigens in vaccines were discussed with an emphasis on assessing their potential for use in veterinary biologics. The meeting also contained presentations on factors to consider when developing *in vitro* assays and on the statistical approaches used to evaluate their results (available at http://www.cfsph.iastate.edu/IICAB/meetings/april2004.htm).

Several interesting new technologies were presented, such as proteomics, Diachemix' differential fluorescence polarisation assays and the Biacore® surface plasmon resonance detection system. CVB is eager to evaluate their potential to augment or replace the animal testing standards of today, with the intent to continue incorporating such 3Rs alternatives into the regulatory framework in the future.

References

USDA-APHIS-VS-CVB (US Department of Agriculture-Animal and Plant Health Inspection Service-Veterinary Services-Center for Veterinary Biologics). "Background and Summary of Activities." http://www.aphis.usda.gov/vs/cvb/lab/background.htm

USDA-APHIS-VS-CVB (US Department of Agriculture-Animal and Plant Health Inspection Service-Veterinary Services-Center for Veterinary Biologics). "CVB Goals for 2003." http://www.aphis.usda.gov/vs/cvb/scenarios/cvb2003goals.pdf

USDA-APHIS (US Department of Agriculture-Animal and Plant Health Inspection Service). 2004. Draft Guideline on Target Animal Safety: Examination of Live Veterinary Vaccines in Target Animals for Absence of Reversion to Virulence (VICH Topic GL41). Federal Register 69(246): 76903-76904 (2004).

USDA-APHIS-VS-CVB (US Department of Agriculture-Animal and Plant Health Inspection Service-Veterinary Services-Center for Veterinary Biologics). Notice No. 04-09 (April 1, 2004) "Use of Humane Endpoints in Animal Testing of Biological Products" http://www.aphis.usda.gov/vs/cvb/notices/2004/04-09.pdf

USDA-APHIS-VS-CVB (US Department of Agriculture-Animal and Plant Health Inspection Service-Veterinary Services-Center for Veterinary Biologics). Notice No. 05-01 (January 13, 2005) "Mouse Safety Testing" http://www.aphis.usda.gov/vs/cvb/notices/2005/01.pdf

Correspondence to

Jodie A. Kulpa-Eddy, D.V.M.
Staff Veterinarian
United States Department of Agriculture (USDA)-
Animal and Plant Health
Inspection Service (APHIS)-Animal Care
4700 River Road, Unit 84
Riverdale, MD 20737-1234
USA
e-mail: jodie.a.kulpa-eddy@aphis.usda.gov

Workshop 5.10
Ecotoxicity – applying the 3Rs

Applying the Three Rs in Acute Ecotoxicity

Argelia Castaño
Environmental Toxicology, CNSA – I.S. Carlos III, Madrid, Spain

Summary
For hazard identification/classification in Environmental Risk Assessment, the first step is to determine the acute ecotoxicity and, on basis of this, to establish a No Observed Effect Concentration (NOEC). Acute ecotoxicity is based on LC50 tests with single species. For a rough estimation of toxicity for classification and labelling purposes killing a large number of vertebrates is not justified. This paper analyses the present situation and solutions proposed to avoid or to reduce vertebrate tests either with more rational testing strategies and/or by developing animal alternatives that can provide data on this first step in the estimation of toxicity.

Keywords: ecotoxicity, three Rs, fish cells, fish embryo, testing strategies

Introduction

The assessment of the adverse effects of chemicals and other stressors on the environment is a difficult task because of the complex network of species and physicochemical factors interrelated in a given ecosystem. Since it is almost impossible to perform tests in complete ecosystems, ecotoxicity testing has to rely on simplified testing systems, but should still be as representative for the whole ecosystem as possible.

The classical approach for Environmental Risk Assessment (ERA) combines calculations of exposure, hazard identification and risk characterisation. Regulatory testing guidelines identify two environmental compartments: terrestrial and aquatic, depending on the exposure scenario.

For hazard identification/classification, the first step is to determine the acute ecotoxicity and then, based on this, the Predicted No Effect Concentration (PNEC) is established. Acute ecotoxicity is based on LC_{50} tests with single species, usually one species per trophic level (aquatic and terrestrial primary producers, invertebrates and vertebrates).

Single species LC_{50} data have been required for more than 30 years in the regulatory guidelines (EC, 1967), although their value is highly questionable with respect to toxicological relevance. To base environmental risk assessment on the death of individuals after short-term exposure to high toxicant concentrations is even more questionable – perhaps with the exception of accidental spills – but is still what we have in hand and what is accepted by the regulatory community.

Today there is a strong political ambition to reduce animal experiments and replace them with alternatives. For a rough estimation of toxicity for classification and labelling purposes it is not justified to kill a large number of vertebrates. Alternatives in ecotoxicology focus on replacement of tests performed with vertebrates. Currently, acute LC_{50} data on vertebrates are required in ERA of chemicals, active substances of plant protection products (PPP), biocides and human and veterinary medicines (EC, 1991; 1992; 1998).

Serious ethical concern has been raised about the use of vertebrates in ERA, since they suffer severe distress and pain when subjected to acutely toxic concentrations of chemicals, which is clearly not acceptable in view of animal welfare. Therefore, it is becoming more and more urgent to find solutions to avoid, or in the worst case to reduce, vertebrate testing, either with more rational testing strategies and/or by developing animal alternatives that can provide data on this first step in estimating toxicity.

Terrestrial vertebrates

Acute oral toxicity data on mammals and birds - as representatives of terrestrial vertebrate wildlife - are required for the registration of active substances (EC, 1991). 3R approaches for mammals in ERA benefit from the process of hazard characterisation for humans (OECD 2001: TG-420, 423, 425). In practice, the elaboration of a testing scheme for wild mammals is generally not necessary, as a great number of laboratory studies with mammals are carried out for the assessment of the risk to humans. Thus, data for mammals on acute, subacute, chronic toxicity and carcinogenicity, as well as findings concerning influences on reproduction and embryonic development are available for each compound. That same data can be used to assess the ecotoxicological risk for free-living mammals. Other than in special circumstances, further toxicological studies on wild mammals would generally not provide additional information that justifies the effort required and the sacrifice of animals involved.

On the other hand and in general, less attention has been paid to finding alternatives for toxicity tests on birds and those developed are limited to reduction and refinement, and refer only to acute tests. The development of new avian toxicity guidelines and updating of existing OECD guidelines was the subject of a joint SETAC/OECD workshop on Avian Toxicity Testing held in December 1994 at Pensacola (USA). Among the recommendations of the OECD working group created after the workshop, a new guideline for avian acute oral toxicity was proposed. (OECD TG-223, October 2002). At present, the guideline is being finalised after the period for public comments ended in February 2003. The new guideline favours determination of only the lethal threshold and non-observed-effect-doses, with a view to reducing the number of animals needed per test. The new guideline provides a sequential testing procedure that optimises the doses and matches the precision of the endpoint with the precision required for hazard assessment and labelling.
(http://www.oecd.org/document/62/0,2340,fr_2649_201185_2348862_1_1_1_1,00.html)

Aquatic vertebrates

While acute toxicity data on mammals and birds are only required for the registration of active substances under EU directive 91/414, acute fish bioassays need to be performed for all cited substances and chemicals, and in some EU countries, for Whole Effluent Assessment.

In the European Union, data requirements for the notification of new substances are listed in the annexes to the Council Directive 67/548/EEC on the approximation of the laws, regulations and administrative provisions relating to the classification, packaging and labelling of dangerous substances (EC, 1967). The quality and amount of data required primarily depends on the amount of the substance to be marketed (EC, 1992). The "base set" data are required for all substances for which the marketing quantity exceeds one ton per year per manufacturer (Annex VII.A of the Directive). The data reported are used to decide on the classification and labelling as well as for hazard and risk assessment of the substance. This "base set" includes acute toxicity for freshwater fish (96 h LC_{50}; OECD TG 203), acute toxicity for daphnids (48 h LC_{50}) and growth inhibition test on freshwater algae (growth rate: 72 h ECr50 and/or biomass: 72 h ECb50). The toxicity value for the most sensitive of these three organisms is used when deciding on classification and labelling as well as for hazard and risk assessment (calculation of Predicted No Effect Concentration – PNEC).

Reduction

The high demand for fish acute data has obviously promoted the interest in finding alternative approaches. Applying the 3Rs philosophy, the biggest advancement in the last years lay in the refinement of acute fish bioassays. The numbers of fish used can be reduced by 60-70% by applying a strategy proposed by Hutchinson and co-workers in 2003 based on the concept of a threshold approach (Hutchinson et al., 2003).

The principle of the threshold approach is based on findings from a number of studies (Weyers et al., 2000; Hutchinson et al., 2003, Jeram et al., 2005) that report that either Daphnia or algae are the most sensitive for approximately 80% of substances tested. Consequently, most of the current risk calculations from base set testing are based on Daphnia and algae data and fish toxicity data are used only for 20% of the substances. For 80% of the tested chemicals the fish LC_{50} test was a wasted effort. In the strategy proposed by Hutchinson et al. (2003), tests on algae and daphnia are performed first, the results are compared, and the lowest value is set as the upper threshold concentration (UTC). Using this concentration, a fish test is carried out. If no toxicity is observed, no further tests are carried out, and the acute fish toxicity result (LC_{50}) is reported as greater than (>) the UTC value. In case the fish show a toxic response, a second test is performed at a lower concentration using a dilution factor of 3.2. The testing continues by stepping down concentrations until no toxicity is observed. The 96h-LC_{50} value can be obtained from all step-down threshold test data. This approach was presented at the 23rd ESAC scientific meeting. A statement was unanimously endorsed on 28 September 2005 supporting the strategy in general. However, ESAC did not approve the step-down procedure, i.e. when fish mortality occurs in the threshold test at the UTC. For the time being, ESAC suggests that a full LC_{50} test should be performed corresponding to the requirements of OECD Test Guideline 203. (http://ecvam.jrc.cec.eu.int/index.htm).

Replacement

Replacing the acute bioassay on adult fish by adopting the fish embryo test (Nagel, 2002) was already proposed and in some countries adopted for WEA.

The rationale for this proposal is based on the animal definition in the Council Directive 86/609/EEC (EC1986; 2003): *'animal' unless otherwise qualified, means any live non-human*

vertebrate, including free-living larval and/or reproducing larval forms, but excluding foetal or embryonic forms".

In 2001, a working group convened by the German Standardisation Organisation (DIN) designed a protocol for an alternative test based on zebrafish embryos as a replacement for the acute fish test in routine whole-effluent testing (DIN, 2001). After an exposure period of 48 h after fertilisation, the toxicological endpoints coagulation, failure to develop somites, lack of tail detachment from the yolk as well as lack of heartbeat are assessed. Since zebrafish only hatch after approx. 72-96 h, this alternative method does not represent an animal test in legal terms. Since the 1st of January 2005, the so-called "zebrafish egg test" has become mandatory in Germany.

The German Federal Environment Agency has submitted a modified protocol for sewage testing to ISO and will submit another modified protocol for chemical testing to the OECD in late 2005, which will not only cover zebrafish, but also other common OECD fish species such as the fathead minnow (Pimephales promelas) and the Japanese medaka (Oryzias latipes). The proposal has been submitted to ECVAM and is on the way to be peer reviewed by ESAC. Since there is a strong commitment to alternative methods in future chemical testing within the framework of REACH (Registration, Evaluation and Authorisation of Chemicals), the new EU chemicals policy (EC, 2001), the fish embryo test seems a very promising candidate for future toxicological routine testing (Braunbeck et al., 2005; Lange et al., 1995; Nagel, 2002).

The use of embryos instead of adult fish represents a considerable advance, however, the use of fish embryos still requires the maintenance of fish facilities, and for some scientific sectors tests with embryos represent a refinement rather than a full replacement of vertebrate testing.

On the way to fully replacing fish individuals, a cytotoxicity test with fish-derived cells seems to be a realistic alternative.

In vitro assays based on fish cell lines have been used in ecotoxicology for screening, for toxicity ranking of chemicals, chemical mixtures, environmental samples and in Toxicity Identification Evaluations (T.I.E.) during the last 30 years (for review, see Castaño et al., 2003). A number of *in vitro* cytotoxicity assays using fish cells have been developed, the majority of them employing cells from salmonid and cyprinid species. As a general rule, good correlations are found among different cell lines, endpoints, and among different laboratories. In addition, there is a good correlation between the *in vitro* data and the *in vivo* LC_{50} fish data in ranking the toxicity. However, when considering sensitivity, it is clear that *in vitro* cell assays have a markedly lower sensitivity than the *in vivo* fish LC_{50} test. On average, the absolute sensitivity is one or two orders of magnitudes less than that of the acute lethality *in vivo* assays (Segner and Lenz, 1993; Castaño et al., 2003). An advantage is that *in vitro* tests do not carry the risk of indicating false positives, but the problem is that they can indicate false negatives (Castaño et al., 2003). The well recognised good correlation between *in vivo* and *in vitro* fish data favours the use of fish cells, although the lack of sensitivity and the appropriate prediction models preclude their general acceptance for testing purposes.

Regarding the sensitivity we should not forget the wide range of sensitivity found among the *in vivo* fish bioassays. As indicated before, the acute fish test has long been a mandatory component in initial toxicity testing. As a consequence, a considerable body of information has been accumulated for existing chemicals. A closer inspection of existing acute fish LC_{50} data, however, reveals differences in orders of magnitude not only between species, but also for the same species between laboratories (e.g. Juhnke and Lüdemann, 1978). On the other hand, and comparing the fish embryo test and cytotoxicity test, for some chemicals there is some evidence that the fish embryo test is slightly more sensitive and more accurately mirrors toxicity to fish than the cytotoxicity tests (Lange et al., 1995; Nagel, 2002). However, it should be noted that for particular substances cytotoxicity tests may be more sensitive than fish embryo tests (e.g., Zabel and Peterson, 1996).

Recent initiatives suggest an approach combining the 3Rs, i.e. reducing using the threshold approach, replacing (refining) by means of a fish-embryo test and replacing with fish cells (Braunbeck et al., in preparation).

Fish *systems* cannot be deleted from the initial "base set" for testing aquatic ecotoxicity, because for between 20 and 40% of the test substances fish represent the most sensitive taxon

Tab. 1: Linear correlations among EC_{50} values from different basal cytotoxicity tests performed in the RTG-2 and R1 fish cell lines and EC_{50} from two *Daphnia magna* assays for a total of 26 chemicals from the Multicenter Evaluation of *In Vitro* Cytotoxicity test (MEIC) study.

MEIC test number	43	44	45	246	247	248
43	1.00	0.99	0.98	0.98	0.96	0.96
44	0.99	1.00	0.99	0.97	0.95	0.96
45	0.98	0.99	1.00	0.97	0.96	0.96
246	0.98	0.97	0.97	1.00	0.99	0.99
247	0.96	0.95	0.96	0.99	1.00	1.00
248	0.96	0.96	0.96	0.99	1.00	1.00
53	0.86	0.85	0.84	0.88	0.90	0.90
54	0.87	0.85	0.85	0.87	0.90	0.90

Test numbers: 43, 44 and 45 = basal cytotoxicity in the R1 fish cell line; 246, 247 and 248 = basal cytotoxicity in RTG-2 fish cell line, 53 and 54 = *Daphnia magna* immobilisation assays (data taken from Clemedson et al., 1998).

(Hutchinson et al., 2003; Jeram et al., 2005). The question is how to recognise those chemicals without performing an *in vivo* fish test. The easiest way is to perform, as a first step, a cytotoxicity test with fish cell lines in order to account for potentially fish-specific effects.

Comparisons between the Daphnia test and fish cytotoxicity test have shown good correlations but differences in sensitivity for both chemicals (see tab. 1, from Clemedson et al 1998) and complex mixtures (see tab. 2, fig. 1). While correlations among EC_{50} from different fish cell lines, cytotoxicity endpoints and within different laboratories involved in the MEIC study were always $r \geq 0.96$, correlations in EC_{50} from Daphnia and EC_{50} from fish cell lines laid on average at 0.86. Therefore, *a priori*, fish cells could be able to recognise toxic compounds that Daphnia does not recognise and *vice versa*. However at the present we cannot determine whether those are fish-specific toxic chemicals, because most chemicals used in the MEIC study were not environmentally relevant and it is difficult to find *in vivo* LC_{50} data on fish. As soon as a fish cytotoxicity test is fully validated and the correction or prediction factor established, it can readily be used as a truly alternative method to identify those substances for which fish are likely to be more sensitive than daphnids or algae.

In cases when there is uncertainty whether fish cytotoxicity can be predicted accurately, the fish test LC_{50} (OECD TG-203) or, if properly validated, the fish embryo test could be applied to make an *in vivo* comparison. Alternatively, the cytotoxicity EC_{50} value (after being corrected by the appropriate prediction model) could indicate whether the sensitivity of fish systems is lower than that of algae or daphnids. This "negative result" indicates that the test substance is not specifically toxic to fish, and no further testing would be required, since subsequent risk assessment procedures would be based on Daphnia or algae EC_{50} values. The proposal ensures that any testing with adult fish will be reduced to the minimum and if embryo fish were fully validated, testing with adult fish would be avoided, and it reduces the number of tests to be carried out with fish embryos to the absolute minimum.

Tab. 2: Comparison between EC_{50} from *D. magna* inmobilisation test and EC_{50} from Neutral Red Uptake on the RTG-2 fish cell line after testing 49 samples of complex mixtures (industrial effluents). 92% of the effluents are toxic for both test. 2/3 correlates positively (r=0.62, p<0.001) (fig 1). 1/3 do not correlates at all. The conclusion is that complementary information was provided by each test (Castaño unpublished data).

Effluents number	*D. magna* LC_{50} 48 h	RTG-2 Cytotoxicity EC_{50} 48 h
1	No Toxicity	Toxicity
2	Toxicity	No Toxicity
41	Toxicity	Toxicity
5	No Toxicity	No Toxicity

References

Braunbeck, T., Böttcher, M., Hollert, H. et al. (2005). Towards an alternative for the acute fish LC_{50} test in chemical assessment: The fish embryo toxicity test goes multi-species – an update. *ALTEX 22, (Linz 2004)*, 87-102.

Braunbeck, T., Castaño, A., Douben, P. et al. (2006). A proposed combined 3 R-based alternatives approach for the replacement of fish LC_{50} testing (in preparation).

Castano, A., Bols, N., Braunbeck, T. et al. (2003). The use of fish cells in ecotoxicology - the report and recommendations of ECVAM workshop 47. *ATLA 31*, 317-351.

Clemedson, C., Andersson, M., Aoki, Y. et al. (1998). In vitro results from 67 toxicity assays used to test reference chemicals 31-50 and a comparative cytotoxicity analysis. *ATLA 26 (Suppl. 1)*, 131-183.

DIN (2001). German standard methods for the examination of water, waste water and sludge – Subanimal testing (group T) – Part 6: Toxicity to fish. Determination of the non-acute-poisonous effect of waste water to fish eggs by dilution limits (T 6). DIN 38415-6, 2001; German Standardisation Organisation.

EC (2003)Directive 2003/65/EC of the European Parliament and of the Council of 22 July 2003 amending Council Directive 86/609/EEC on the approximation of laws, regulations and administrative provisions of the Member States regarding the protection of animals used for experimental and other scientific purposes .

EC (1967). Council Directive 67/548/EEC of 18 August 1967 on the approximation of the laws, regulations and administrative provisions relating to the classification, packaging and labelling of dangerous substances. O*fficial Journal of European Communities*, L96/1. Office for Official Publications of the European Communities. Luxembourg.

EC (1986). Council Directive 86/609/EEC of 24 November 1986 on the approximation of laws, regulations and administrative provisions of the Member States regarding the protection of animals used for experimental and other scientific

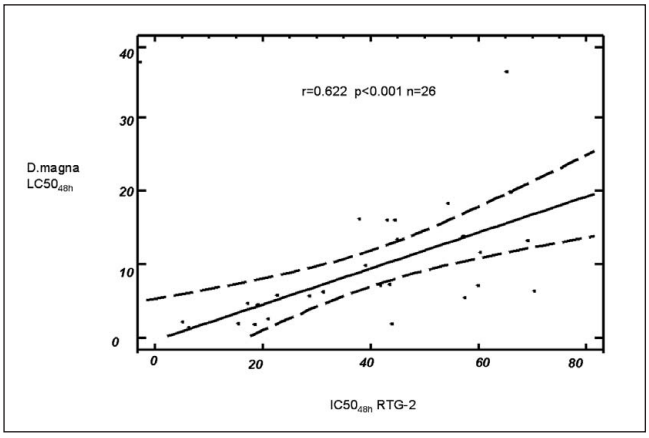


Fig 1: Linear Correlation between EC_{50} from *D. magna* immobilisation test and EC_{50} from Neutral Red Uptake on the RTG-2 fish cell line after testing 26 samples of complex mixtures (industrial effluents), Castaño (unpublished data).

purposes. *Official Journal of the European Community* L358/1, 1-28. Office for Official Publications of the European Communities. Luxembourg.

EC (1991) Council Directive 91/414/EEC of 15 July 1991 concerning the placing of plant protection products on the market. *Official Journal of European Communities* L 230. Office for Official Publications of the European Communities, Luxembourg.

EC (1992). Council Directive 92/32/EEC of 5 June 1992 amending for the 7th time Council Directive 67/548/EEC. *Official Journal of European Communities*, L 154. Office for Official Publications of the European Communities, Luxembourg.

EC (1998). Council Directive 98/8/EC of 16 February 1998 concerning the placing of biocidal products on the market. *Official Journal of European Communities*, L123/1. Office for Official Publications of the European Communities. Luxembourg.

EU COM (2001). *Commission of the European Communities: White paper, strategy for a future chemicals policy*. Brussels, Belgium.

Hutchinson, T. H., Barrett, S., Busby, M. et al. (2003). A strategy to reduce the numbers of fish used in acute ecotoxicity testing of pharmaceuticals. *Environmental Toxicology and Chemistry 22*, 60-65.

Jeram, S., Riego Sintes, J. M., Halder, M. et al. (2005). A strategy to reduce the use of fish in acute ecotoxicity testing of new chemical substances notified in the European Union. *Regulatory Toxicology and Pharmacology, 42(2)*, 218-224.

Juhnke, I. and Lüdemann, D. (1978).Ergebnisse der Untersuchung von 200 chemischen Verbindungen auf acute Fischtoxizität mit dem Goldorfentest. *Z. Wasser, Abwasser-Forsch. 11*, 161-164.

Lange, M., Gebauer, W., Markl, J. and Nagel, R. (1995). Comparison of testing acute toxicity on embryo of zebrafish, Brachydanio rerio and RGT-2 cytotoxicity as possible alternatives to the acute fish test. *Chemosphere 30*, 2087-2102.

Nagel, R. (2002). DarT: The embryo test with the zebrafish Danio rerio – a general model in ecotoxicology and toxicology. *ALTEX 19*, 38-48.

OECD Guidelines for the Testing of Chemicals (1992). Section 2: Effects on Biotic Systems Test No. 203: Acute Toxicity for Fish. Paris, France: Organization for Economic Cooperation and Development.

OECD Guidelines for the Testing of Chemicals (1992). Section 2: Effects on Biotic Systems Test No. 223 Avian Acute Oral Toxicity Proposal A PDF and Proposal B PDF, Deadline for public comments passed: February 2003.

OECD Guidelines for the Testing of Chemicals (2001). Section 4: Health Effects Test No. 420: Acute Oral Toxicity - Fixed Dose Procedure. Paris, France: Organization for Economic Cooperation and Development.

OECD Guidelines for the Testing of Chemicals (2001). Section 4: Health Effects Test No. 423: Acute Oral toxicity - Acute Toxic Class Method. Paris, France: Organization for Economic Cooperation and Development.

OECD Guidelines for the Testing of Chemicals (2001). Section 4: Health Effects Test No. 425: Acute Oral Toxicity: Up-and-Down Procedure. Paris, France: Organization for Economic Cooperation and Development.

Segner, H. & Lenz, D. (1993). Cytotoxicity assays with the rainbow trout R1 cell line. *Toxicology in vitro 7*, 537-540.

Weyers, A., Sokull-Klüttgen, B., Baraibar-Fentanes and J., Vollmer, G., 2000. Acute toxicity data: a comprehensive comparison of results of fish, Daphnia and algae tests with new substances notified in the EU. *Environ. Toxicol. Chem. 19*, 1931-1933.

Zabel, E. W. and Peterson, R. E. (1996) TCDD-like activity of 2,3,6,7-tetrachloroxanthene in rainbow trout early life stages and in a rainbow trout gonadal cell line (RTG-2). *Environ. Toxicol. Chem. 15*, 2305-2309.

Acknowledgment

This work was partially financed by Spanish project: BIO-02-1912.

Correspondence to

Dr. Argelia Castaño
Environmental Toxicology
CNSA – I. S. Carlos III
Cra Majadahonda
Madrid, Spain
e-mail: castano@isciii.es

Ecotoxicological Tests in Non-Ecotoxicological Research: Contribution to the Three Rs

Use of luminescent photobacteria for evaluating the toxicity of 47 MEIC reference chemicals

Anne Kahru
National Institute of Chemical Physics and Biophysics, Estonia

Summary
The luminescent bacterium Photobacterium phosphoreum was used for toxicity testing of 47 MEIC reference chemicals. Five-minute EC_{50} values (the concentration of chemical, in mM, which reduces the light output of the bacteria by 50% after contact for 5 minutes) were calculated and correlated with data from the literature: octanol/water partition coefficients, acute $L(E)C_{50}$ data from the Microtox™ and Daphnia tests, acute in vitro toxicity data for fish, animal and human cell lines as well as in vivo data for rodents, dog and man. The log-log correlation coefficients (R^2) ranged between 0.20-0.79. Thus, the application of a photobacteria-based toxicity test (as part of an in vitro test battery) at the screening stage of all areas of in vitro toxicological research (not only in ecotoxicological studies) will contribute to the Three Rs concept.

Keywords: Photobacterium phosphoreum, Microtox™, MEIC chemicals, rodent and human acute lethal toxicity, three Rs

Introduction

Over the last twenty years, alternative, non-animal test systems (mainly eukaryotic cell cultures) have been introduced to supplement and, in some cases, to replace toxicity tests using animals (Carere et al., 2002), contributing to the Three R's concept (replacement, reduction, refinement) introduced by Russel & Burch in 1959.

Several international projects aimed to validate the predictability of *in vitro* tests have been conducted, e.g. the MEIC (**M**ulticenter **E**valuation of ***I**n Vitro* **C**ytotoxicity) programme (1989-1996), initiated by Björn Ekwall and organised by the Scandinavian Society of Cell Toxicology (Bondesson et al., 1989; Ekwall et al., 1989). The MEIC programme was established to evaluate the relevance and reliability of a wide variety of *in vitro* tests for predicting general toxicity in more complex biologic systems. The final goal of the MEIC project was to define optimal batteries of *in vitro* toxicity tests suitable for replacing or supplementing animal tests in predicting the toxicity of chemicals in humans. The rationale behind that approach lay in the fact that toxicity data concerning effects on basic cellular functions and/or structures have a good correlation with *in vivo* data concerning human lethal toxicity (Ekwall, 1999).

Approximately 60 *in vitro* cytotoxicity assays (using human/animal cells and other bioassay endpoints) were tested on 50 compounds of human toxicological interest. MEIC efforts resulted in a series of eight papers in *ATLA* compiling *in vitro* assay, human and rodent toxicity data and comparing their responses (for references, see Ekwall et al., 2000).

In the current work, the reconstituted, freeze-dried (viable) bacteria *Photobacterium phosphoreum* were used for the short-term toxicity testing of 47 chemicals from the MEIC list. Three MEIC reference chemicals were not analysed: dextropropoxyphene chloride (no. 22), thallium sulphate (no. 30) and amphetamine sulphate (no. 47).

P. phosphoreum are naturally luminescent marine bacteria that emit light as a result of their metabolism. Damage to the central cellular metabolism caused by a toxic substance results in a decrease in light output by the bacteria (reduction in light output represents physiological inhibition, not just mortality), the light loss being proportional to the degree of toxicity (Hastings, 1978). In this work the concentration-effect curves for 47 MEIC chemicals were measured and five-minute EC_{50} values (the concentration of chemical which reduces the light output of the bacteria by 50% after contact for 5 minutes) were calculated. These EC_{50} data from the photobacterial test were correlated with data from the literature: octanol/water partition coefficients, $L(E)C_{50}$ data from the Microtox™ and *Daphnia* tests, and acute toxicity data for freshly isolated fish hepatocytes, animal and human cell lines, rodents, dog and man. This work is an update of our previous study (Kahru and Borchardt, 1994) on the toxicity of 39 MEIC chemicals in *P. phosphoreum*.

Materials and methods

Origin of chemicals

MEIC reference chemicals tested are listed in table 1. Chemicals No. 1, 6, 13, 21, 23, 24, 25, 29, 36, 40, 41, 42, 49, 50 were from Sigma; chemicals No. 2, 3, 7, 12, 14, 18, 19, 20, 26, 27, 28, 46 where from Merck; chemical No. 15 (95% purity) was from Kemira (Finland); chemicals No. 5, 17, 31, 34, 37, 38, 43, 44, 48

were from Aldrich; Chemical No. 4 (in the form of pharmacological tablets) was from Rivofarm (Switzerland); chemical No. 32 was from Dr. Schuchardt (Germany); chemical No. 45 was from Boehringer-Mannheim; chemical No. 33 was from Carlo Erba, and chemicals 8, 9 and 10 (HPLC-grade) were from Russia.

Solvents

Chemicals No. 3, 5, 23, 25, 26, 27, 28, 31, 35, 37, 38, 41, 42, 45, 46, 49 were tested in 2% NaCl; chemicals No. 7, 8, 9, 10, 12, 13, 14, 18, 19, 20, 21, 48, 50 were tested in phosphate-buffered 2% NaCl (PBNaCl) (pH=7); chemicals No. 4, 11, 15, 17, 24, 29, 33, 34, 36, 39, 40, 43, 44 were tested in 2% NaCl + 1.5% methanol; chemicals No. 1, 2, 16, 32 were tested in PBNaCl + 1.5% methanol, and chemical Nr. 6 was tested in 2% NaCl + 4% DMSO. The chemicals were handled and the toxicity testing was performed essentially as described in our previous papers (Kahru, 1993; Kahru and Borchardt, 1994).

Test bacteria

Laboratorially prepared freeze-dried photobacterial reagent (*P. phosphoreum* strain FEI 162095 registered in the Finnish Environment Institute) was used. The freeze-dried bacteria were reconstituted by adding 20 ml 2% NaCl into the reagent bottle and incubating it at 4ºC for 20 minutes before use. The average number of bacteria used per assay was about one million.

Statistical analysis

Comparisons between the data sets were made by calculating linear regression using the log-transformed toxicity data. A perfect fit of data where both the photobacterial and other *in vitro* or *in vivo* results were identical would produce an R^2 of 1.0, a slope of 1.0 and a Y-intercept of 0, and the proximity of the calculated values to these absolutes is indicative of the predictive capacity of the X-variable (photobacterial toxicity data) for the Y-variable (data used for the comparison).

Results and discussion

The toxicity of 47 MEIC chemicals was determined using the inhibition of light output of naturally luminescent *Photobacterium phosphoreum* as a toxicity endpoint. The five-minute EC_{50} values, calculated both as mg/L and mM, are presented in table 1. The EC_{50} data for 39 MEIC chemicals obtained in our earlier study (Kahru and Borchardt, 1994) were updated with the toxicity data for paraquat (No. 25), arsenic trioxide (No. 26), warfarin (No. 31), barium nitrate (No. 37), hexachlorophene (No. 38), orphenadrine hydrochloride (No. 42), diphenylhydantoine (No. 44) and sodium oxalate (No. 46) (tab. 1).

Table 1 shows that the five least toxic MEIC chemicals for *P. phosphoreum* (arranged in increasing toxicity, in mM) were potassium chloride (No. 50, EC_{50}=2830 mM), ethylene glycol (No. 7, EC_{50}=2674 mM), methanol (No. 8, EC_{50}=786.5 mM), ethanol (No. 9, EC_{50}=429.8 mM) and sodium chloride (No. 13, EC_{50}=292.6 mM) (tab. 1), whereas the five most toxic MEIC chemicals (arranged in decreasing toxicity, mM) were mercuric

Tab. 1: Five-minute EC_{50} values of the 47 MEIC chemicals determined using the luminescent photobacteria (*Photobacterium phosphoreum*) luminescence inhibition assay

No.	MEIC Chemical	5-minute EC_{50} [mg/L]	5-minute EC_{50} [mM]
1	Paracetamol	2320.00	15.34
2	Acetylsalicylic acid	1140.00	6.33
3	Iron (II) sulphate	2340.00	8.42
4	Diazepam	214.6	0.75
5	Amitriptyline hydrochloride	10.26	0.03
6	Digoxin	1130.00	1.45
7	Ethylene glycol	166000.00	2674.40
8	Methanol	25200.00	786.52
9	Ethanol	19800.00	429.78
10	Isopropanol	11200.00	186.39
11	1,1,1-Trichloroethane	264.70	1.98
12	Phenol	134.00	1.42
13	Sodium chloride	17100.00	292.61
14	Sodium fluoride	9870.00	235.06
15	Malathion	298.73	0.90
16	2,4-Dichloro-phenoxyacetic acid	128.00	0.58
17	Xylene	102.00	0.96
18	Nicotine	434.00	2.68
19	Potassium cyanide	1730.00	26.57
20	Lithium sulphate	30000.00	234.47
21	Theophylline	2150.00	11.93
23	Propranolol hydrochloride	97.54	0.33
24	Phenobarbital	141.19	0.61
25	Paraquat (Methyl viologen)[a]	14830.00	57.66
26	Arsenic trioxide[a]	7.58	0.04
27	Copper (II) sulphate	138.10	0.55
28	Mercury (II) chloride	0.16	0.0006
29	Thioridazine hydrochloride	7.76	0.02
31	Warfarin[a]	90.35	0.29
32	Lindane	60.82	0.21
33	Chloroform	658.10	5.51
34	Carbon tetrachloride	563.30	3.66
35	Isoniazid	4220.00	30.78
36	Dichloromethane	1850.00	21.78
37	Barium nitrate[a]	29400.00	112.49
38	Hexachlorophene[a]	45.89	0.11
39	Pentachlorophenol	10.33	0.04
40	Verapamil hydrochloride	439.30	0.89
41	Chloroquine phosphate	91.60	0.18
42	Orphenadrine hydrochloride[a]	119.40	0.39
43	Quinidine sulphate	88.14	0.12
44	Diphenylhydantoin[a]	21.51	0.09
45	Chloramphenicol	205.00	0.63
46	Sodium oxalate[a]	31500.00	235.07
48	Caffeine	2150.00	11.07
49	Atropine sulphate	3092.00	4.57
50	Potassium chloride	150000.00	28380.00

[a] chemicals that were not analysed for toxicity in our previous paper (Kahru and Borchardt, 1994)

chloride (No. 28, EC_{50}=0.0006 mM), thioridazine hydrochloride (No. 29, EC_{50}=0.02 mM), amitriptyline hydrochloride (No. 5, EC_{50}=0.03 mM), arsenic trioxide (No. 26, EC_{50}=0.04 mM) and pentachlorophenol (No. 39, EC_{50}=0.05 mM).

The *P. phosphoreum* toxicity data for 47 MEIC chemicals were compared with the toxicities of the corresponding chemicals in other test systems (data from the literature) by calculating linear regression using the log-transformed toxicity data. The test systems with respective toxicity endpoints (including the exposure time used for the respective test), the number of data pairs used for the correlation and the equation for the log-log regression line are presented in table 2.

Correlation with lipophilicity (owp)

The octanol water partition (owp) coefficient characterises the lipophilicity of a substance. As a general rule, substances with a higher lipophilicity are more toxic to living cells, mainly due to their enhanced ability to cross the plasma membrane and enter the cell. Table 2 shows that our data are in full agreement with this theory: a relatively high correlation (R^2=0.62) was obtained

Tab. 2: Correlation between EC_{50} values determined using *Photobacterium phosphoreum* (our data) and toxicity data for matching MEIC chemicals determined using other test systems (data from the literature).

Test system	Cytotoxicity endpoint	R^2 log-log	No. of data pairs correlated	Equation for the log-log regression line	Ref.
Physico-chemical properties of a substance					
Physico-chemical properties of a substance	Octanol/water partition coefficient, owp	0.62	27	y = -1.26x + 2.22	Halle et al., 1991
Ecotoxicological tests					
Vibrio fischeri (Microtox™)	Inhibition of light output, 5- or 15-minute EC_{50}	0.64	42	y = 0.86x - 0.07 Fig. 1A	Calleja et al., 1993
Daphnia magna	Immobilisation, 24-h LC_{50}	0.69	38	y = 0.98x - 1.09 Fig. 1B	Lilius et al., 1994
Rainbow trout hepatocytes (freshly isolated)	Plasma membrane integrity, measured by ^{86}Rb-leakage, 3-h EC_{50}	0.53	40	y = 0.79x + 0.63 Fig. 1C	Lilius et al., 1994
***In vitro* assays with animal or human cell lines**					
Rat hepatocytes	Average of three endpoints (morphometric changes, % of viable cells, lactate dehydrogenase release), 24-h CT_{50}	0.66	46	y = 0.73x - 0.34 Fig. 2A	Shrivastava et al., 1992
Chick forebrain primary neuron cultures	Integrity of mitochondria, assayed by tetrazolium MTT, 3-h EC_{50}	0.59	46	y = 0.63x + 0.26	Weiss et al., 1993[a]
Chick forebrain primary neuron cultures	Lysosomal integrity, analysed by neutral red uptake (NRU), 4-h EC_{50}	0.55	47	y = 0.69x + 0.01	Weiss et al., 1993[a]
Various animal or human cell lines	Geometric mean of two or more IC_{50} values per substance. Various endpoints, CT_{50}x	0.79	27	y = 0.83x - 0.22 Fig. 2B	Halle et al., 1991
***In vivo* data**					
Rat	Acute oral LD_{50}	0.37	43	y = 0.45x + 0.10	Weiss et al., 1993[a]
Rat	Acute i.p. LD_{50}	0.29	36	y = 0.38x - 0.14	Weiss et al., 1993[a]
Rat & mice	Acute oral LD_{50}	0.42	47	y = 0.42x + 0.14 Fig. 3A	Shrivastava et al., 1992
Mouse	Acute oral LD_{50}	0.41	42	y = 0.45x - 0.11	Weiss et al., 1993[a]
Mouse	Acute i.p. LD_{50}	0.41	42	y = 0.45x - 0.27	Weiss et al., 1993[a]
Dog	acute lethal i.v. dose, LDL0	0.50	20	y = 0.67x - 0.95	Weiss et al., 1993[a]
Human	Acute lethal plasma concentration, LC	0.32	46	y = 0.52x - 0.96 Fig. 3B	Barile et al., 1994

[a] reported by RTECS (Registry of Toxic Effects of Chemical Substances)

when the toxicity of MEIC chemicals to *P. phosphoreum* was compared to the owp of these chemicals (owp values were taken from Halle et al., 1991).

Correlation with ecotoxicological tests

When the *P. phosphoreum* data were compared with Microtox™ data (from Calleja et al., 1993), the log-log R^2 value was 0.64 (tab. 2; fig. 1A). The high correlation is not surprising, as the test used in the current study and the Microtox™ test are both based on luminescent photobacteria. Moreover, the strain used in the Microtox™ test was formerly also called *Photobacterium phosphoreum* (now *Vibrio fischeri*). We have previously compared the *P. phosphoreum* data (i.e. the strain used by us) with *V. fischeri* (Microtox™) data and revealed a good correlation (log-log correlation coefficient R^2 for 25 laboratory chemicals was 0.94; Kahru, 1993).

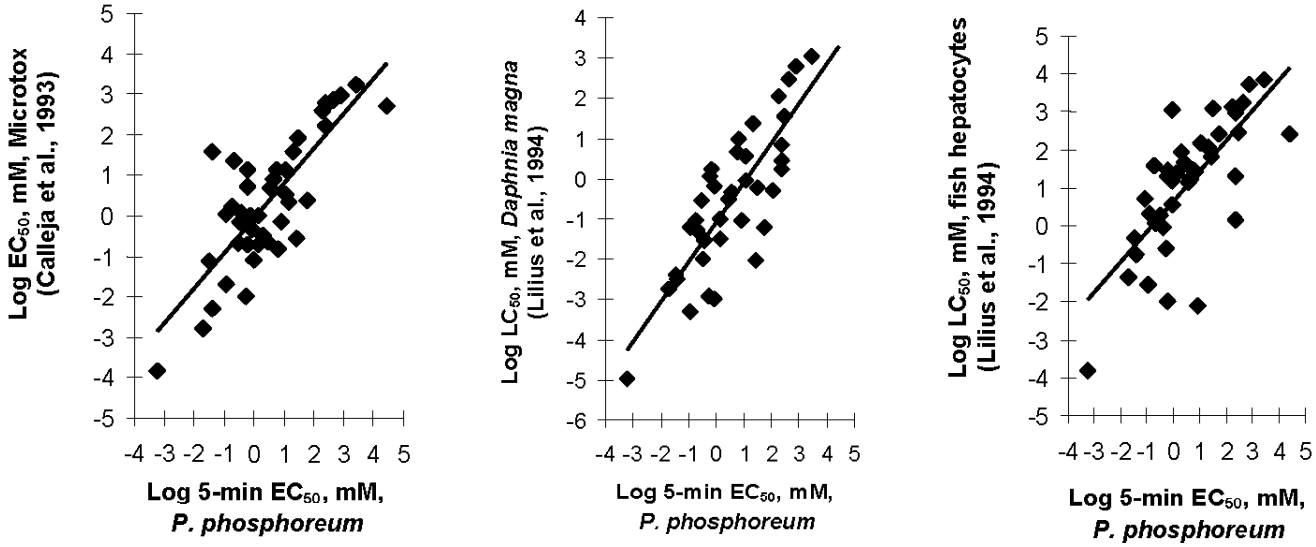
Table 2 and figure 1B show that there was also a relatively high log-log correlation (R^2=0.69) between *P. phosphoreum* toxicity data and that of the *Daphnia* test, a widely used toxicity test in aquatic toxicology, where the immobilisation of aquatic invertebrate *Daphnia magna* after 24 h incubation with chemicals was used as the toxicity endpoint (Lilius et al., 1994). However, the toxicity of most of the chemicals to *Daphnia* exceeded the toxicity to photobacteria by two orders of magnitude. One possible explanation for different sensitivities could be the difference in exposition times (24 hours in the case of *Daphnia versus* 5 min in the case of photobacteria) (tab. 2), i.e. the short exposure to the toxicant may be too short for some types of cellular damages to occur (e.g., caused by bioaccumulation of a toxicant). Correlation of the *P. phosphoreum* toxicity assay with freshly isolated rainbow trout hepatocytes (data taken from Lilius et al., 1994) was reasonably good (log-log R^2=0.53), whereas photobacteria were about 40-fold more sensitive (fig. 1C). It should be mentioned that fish cell lines as well as primary cell cultures are of growing importance in risk assessment of chemicals as substitutes for the acute lethality test in fish usually performed according to OECD guidelines (Fent, 2001; Castaño et al., 2003).

Correlation with cell line assays

Table 2 and figure 2B show that the highest log-log correlation (R^2=0.79) was obtained when *P. phosphoreum* data were compared with the average *in vitro* basal cytoxicity (IC_{50}x, the geometric mean of two or more IC_{50} values per substance, determined by different investigators using various animal or human cell lines and different cytotoxicity endpoints, compiled by Halle et al. (1991). Moreover, the average EC_{50}/IC_{50}x value was 4.3, indicating that the average sensitivity of the photobacterial assay was comparable with that of eukaryotic cell cultures, being at the same time a much less expensive and less laborious toxicity test.

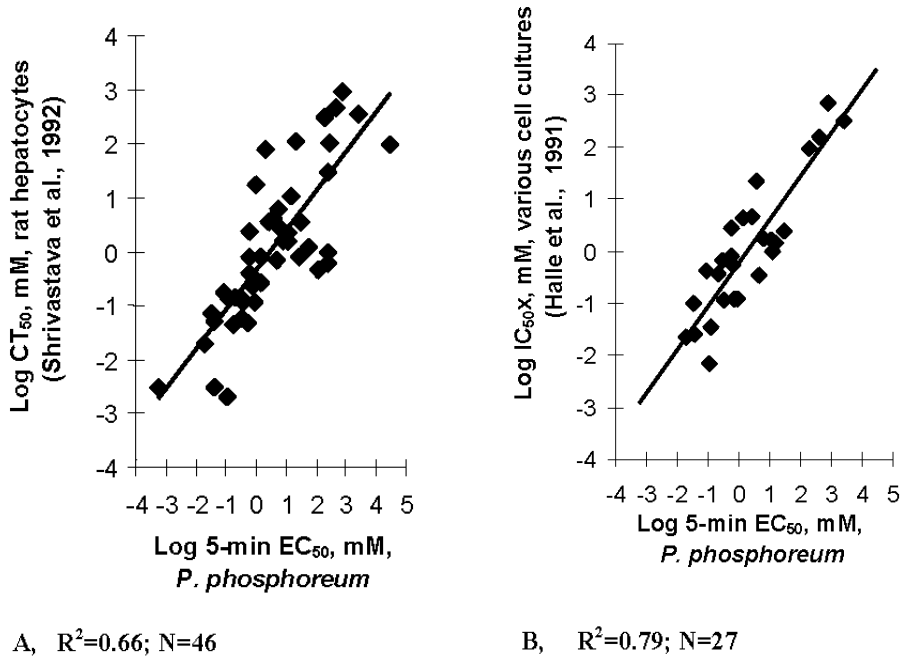
Correlation with *in vivo* data for rodents, dog and human

Table 2 and figure 3 present the comparison of the toxicity of MEIC chemicals in *P. phosphoreum* with *in vivo* toxicity data for the corresponding chemicals obtained from the literature (acute oral LD_{50} and i.p. LD_{50} values for rats and mice, acute i.v. LD values for dog and acute oral and blood plasma LD values for man were used for the correlation, see references in tab. 2). It can be seen that the correlation of five-minute EC_{50} data with the acute oral LD_{50} data for rodents (tab. 2 and fig. 3A) yielded very



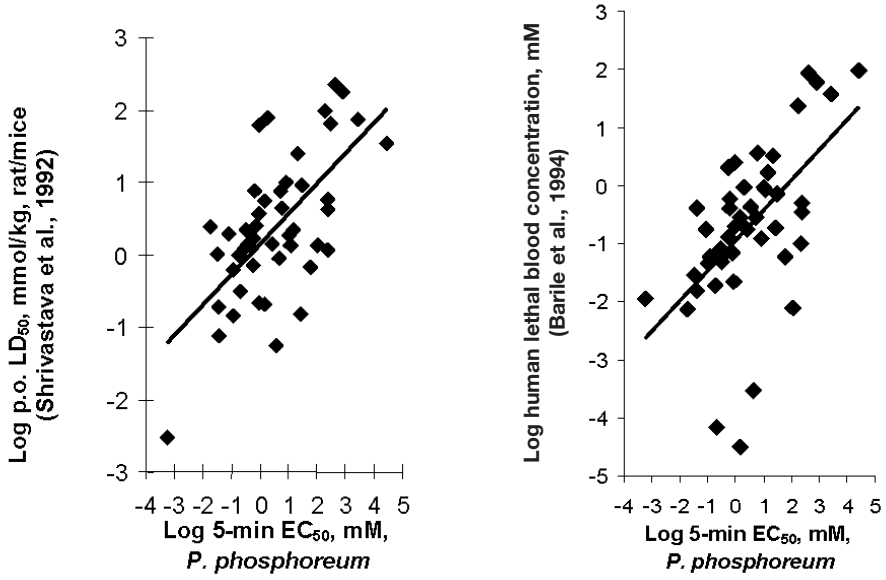
A, R^2=0.64; N=42 B, R^2=0.69; N=38 C, R^2=0.53; N=40

Fig. 1: Logarithms of EC_{50} values (mM) obtained using the *Photobacterium phosphoreum* test *versus* logarithms of $L(E)C_{50}$ values (mM) obtained using the Microtox test for 42 MEIC chemicals (A), *Daphnia magna* test for 38 MEIC chemicals (B) and freshly isolated rainbow trout hepatocyte assay for 40 MEIC chemicals (C). Origin of data used for the comparison and information on the tests are presented in table 2.



A, R^2=0.66; N=46 B, R^2=0.79; N=27

Fig. 2: Logarithms of EC_{50} values (mM) obtained using the *Photobacterium phosphoreum* test *versus* logarithms of CT_{50} values (mM) obtained using rat hepatocyte cultures for 46 MEIC chemicals (A), and logarithms of IC_{50}x values (mM) obtained on various animal or human cell lines for 27 MEIC chemicals (B). Origin of data used for the comparison and information on the tests are presented in table 2.



A, R^2=0.42; N=47 B, R^2=0.32; N=46

Fig. 3: Logarithms of EC_{50} values (mM) obtained using the *Photobacterium phosphoreum* test *versus* logarithms of acute rat/mice p.o. concentrations (mmol/kg) for 47 MEIC chemicals (A) and logarithms of acute lethal blood plasma concentration, LD; mM) in humans for 46 MEIC chemicals (B). Origin of data used for the comparison is presented in table 2.

similar correlation coefficients (R^2= 0.29-0.42) for all rodent data sets, whereas the R^2 values were slightly higher if mouse data were used for the comparison. The highest correlation (R^2=0.50) was obtained, when *P. phosphoreum* data were correlated with dog lethal i.v. values (referenced in Weiss and Sayer, 1993) (tab. 2).

The correlation of photobacterial data with human data (tab. 2 and fig. 3B) was not impressive (R^2= 0.20-0.32). However, the predictive potential of the chick forebrain primary neuron culture MTT (see also tab. 2) EC_{50} data for the determination of toxic doses in humans (R^2=0.29; Weiss and Sayer, 1993) was comparable to that of *P. phosphoreum*. Also, the predictive

potential of rat i.p. LD_{50} data for man oral LD (R^2=0.19; Weiss and Sayer, 1993) was comparable to that of *P. phosphoreum*. It must also be stressed, that, due to ethical necessity, estimations of human chemical toxicity are based on ill-defined populations of people and, often, only estimates of dosage. This could be one of the reasons for the relatively poor correlation observed.

Conclusions

This paper is an update of our previous study in which the toxicity of 39 MEIC chemicals was studied using naturally luminescent marine bacteria, *Photobacterium phosphoreum*, as test organisms (Kahru and Borchardt, 1994). In the current study, that data set was updated with the toxicity data for an additional 8 MEIC chemicals (paraquat, arsenic trioxide, warfarin, barium nitrate, hexachlorophene, orphenadrine hydrochloride, diphenylhydantoine and sodium oxalate) (tab. 2) and correlated with data from the literature: octanol/water partition coefficients, $L(E)C_{50}$ data from the Microtox™ test and *Daphnia*, and acute toxicity data for freshly isolated fish hepatocytes, animal and human cell lines, rodents, dog and man.

Our data showed that luminescent bacteria (e.g., *P. phosphoreum*, *Vibrio fischeri*) could be used at the screening level to test the safety of chemicals, as this assay system is very cheap, easy to use, well reproducible, very rapid, and could be more easily automated than conventional testing systems (e.g. cell cultures, experimental animals). These photobacterial tests also yield very similar results (this paper, Kahru, 1993; Kahru et al., 1996 and Jennings et al., 2001). Moreover, the *V. fischeri* NRRL B-111 77 test is commercially available under several trade marks (ToxAlert™, Microtox™, LUMIStox™, BioTox™), and photobacterial toxicity data exist for more than 1000 chemicals (Kaiser and Devillers, 1994). Therefore, the application of photobacteria-based toxicity tests (as part of an *in vitro* test battery) to all areas of *in vitro* toxicological research, not only ecotoxicological studies (Bispo et al., 1999; Kahru et al., 2000; Castillo et al., 2001; Loibner et al., 2004), should be seriously considered, as this could save a lot of money, manpower and lives of experimental animals and thus contribute to the Three R's concept (replacement, reduction, refinement) introduced by Russel & Burch in 1959.

References

Barile, F. A., Dierickx, P. J. and Kristen, U. (1994). In vitro cytotoxicy testing for prediction of acute human toxicity. *Cell Biology and Toxicology 10*, 155-162.

Bispo, A., Jourdain, M. J.and Jauzein, M. (1999). Toxicity and genotoxicity of industrial soils polluted by polycyclic aromatic hydrocarbons (PAHs). *Organic Geochemistry 30*, 947-952.

Bondesson, I., Ekwall, B., Hellberg, S. et al. (1989) MEIC: a new international multicenter project to evaluate the relevance to human toxicity of in vitro cytotoxicity tests. *Cell Biology and Toxicology 5*, 331-348.

Calleja. M. C., Persoone, G. and Geladi, P. (1993). The predictive potential of a battery of ecotoxicological tests for human acute toxicity, as evaluated with the first 50 MEIC chemicals. *ATLA 21*, 330-349.

Carere, A., Stammati, A. and Zucco, F. (2002). In vitro toxicology methods: impact on regulation from technical and scientific advancements. *Toxicology Letters 127*, 153–160.

Castaño, A., Bols, N., Braunbeck, T. et al. (2003).The use of fish cells in ecotoxicology. *ATLA 31*, 317-351.

Castillo, M., Alonso, M. C., Riua, J. et al. (2001). Identification of cytotoxic compounds in European wastewaters during a field experiment. *Analytica Chimica Acta 426*, 265–277.

Ekwall, B., Bondesson, S., Castell, J. V. et al. (1989). Cytotoxicity evaluation of the first ten MEIC chemicals: acute lethal toxicity in man predicted by cytotoxicity in five cellular assays and by oral LD50 in rodents. *ATLA 17*, 83-100.

Ekwall, B. (1999). Overview of the final MEIC results: II. The in vitro–in vivo evaluation, including the selection of a practical battery of cell tests for prediction of acute lethal blood concentrations in humans. *Toxicology in Vitro 13*, 665–673.

Ekwall, B., Ekwall, B. and Sjostrom, M. (2000). MEIC evaluation of acute systemic toxicity. Part VIII. Multivariate partial least squares evaluation, including the selection of a battery of cell line tests with a good prediction of human acute lethal peak blood concentrations for 50 chemicals. *ATLA 28*, 201–234.

Fent, K. (2001) Fish cell lines as versatile tools in ecotoxicology: assessment of cytotoxicity, cytochrome P4501A induction potential and estrogenic activity of chemicals and environmental samples. *Toxicology in Vitro 15*, 477-488.

Halle, W., Baeger, I., Ekwall, B. and Spielmann, H. (1991). Correlation between in vitro cytotoxicity and octanol/water partition coefficient of 29 substances from the MEIC programme. *ATLA 19*, 338-343.

Hastings, J. W. (1978). Bacterial bioluminescence: An Overview. In M. A. DeLuca (ed.), *Methods in Enzymology*, 57 (125-135). New York: Academic Press.

Jennings, V. L. K., Rayner-Brandes, M. H. and Bird, D. J. (2001). Assessing chemical toxicity with the bioluminescent photobacterium (Vibrio fischeri): A comparison of three commercial systems, *Water Research 35*, 3448–3456.

Kahru, A. (1993). In vitro toxicity testing using marine luminescent bacteria (Photobacterium phosphoreum): the Biotox™ test. *ATLA 21*, 210-215.

Kahru, A. and Borchardt, B. (1994) Toxicity of 39 MEIC chemicals to bioluminescent photobacteria (the BiotoxTM test): correlation with other tests ystems. *ATLA 22*, 147-160.

Kahru, A., Tomson, K., Pall, T. and Külm, I. (1996). Study of toxicity of pesticides using luminescent bacteria Photobacterium phosphoreum. *Water Science and Technology 33*, 147-154.

Kahru, A., Põllumaa, L., Reiman, R. et al. (2000). The toxicity and biodegradability of eight main phenolic compounds characteristic to the oil-shale industry wastewaters: a test battery approach. *Environmental Toxicology 15*, 431-442.

Kaiser, K. L. E. and Devillers, J. (1994). *Ecotoxicity of Chemicals to Photobacterium phosphoreum*. Vol.2 in series of Handbooks on Ecotoxicity of Chemicals. Yverdon, Switzerland: Gordon and Breach Science Publishers.

Lilius, H, Isomaa, B. and Holström, T. (1994). A comparison oif the toxicity of 50 reference chemicals to freschly isolated rainbow trout hepatocytes and Daphnia magna. *Aquatic Toxicology 30*, 47-60.

Loibner, A. P., Szolar, O. H. J., Braun, R. and Hirmann, D. (2004). Toxicity testing of 16 priority polycyclic aromatic hydrocarbons using Lumistox®. *Environmental Toxicology and Chemistry 23 (3)*, 557-564.

Russel, W. M. S. and Burch, R. L. (1959). *The principles of Humane Experimental Technique*. London: Methuen.

Shrivastava, R., Delomenie, C., Chevalier, A. et al. (1992). Comparison of in vivo acute lethal potency and in vitro cytotoxicity of 48 chemicals. *Cell Biology and Toxicology 8*, 157-170.

Weiss, M. T. and Sawyer, T. W. (1993). Cytotoxicity of the MEIC test chemicals in primary neurone cultures. *Toxicology in Vitro 7 (5)*, 653-667.

Acknowledgments

Part of this work was presented at the EC Biomedical & Health Research Biomed Workshop, Patras-Greece, Nov. 17-18, 1994. This work has been supported by the Estonian Science Foundation. The skillful help of Taisia Rõlova (NICPB, Estonia) in preparation of the manuscript is acknowledged.

Correspondence to

Anne Kahru, PhD
National Institute of Chemical Physics and Biophysics
Estonian Academy of Sciences
Akadeemia tee 23
Tallinn 12618
Estonia
e-mail: anne@kbfi.ee

Application of *In Vitro* Alternative Methods to Ecotoxicology

Shinobu Wakuri[1], Yutaka Matsumoto[2], Makoto Hayashi[3] and Noriho Tanaka[1]

[1]Hatano Research Institute, Food and Drug Safety Center, Hadano, Kanagawa, Japan; [2]Hokkaido Research Institute for Environmental Sciences, Sapporo, Japan; [3]Division of Genetics and Mutagenesis, National Institute of Health Sciences, Tokyo, Japan

Summary

We applied in vitro assays using cultured cells for the ecotoxicological assessment of various samples collected from the environment. The results of cytotoxicity assays on water samples from rivers and cooling towers correlated well with the respective water analyses. However, some samples induced different responses between the two cell lines employed. Extracts of solid samples revealed the order of contaminations in the phototoxicity test and mouse lymphoma mutation assay. Therefore, the in vitro toxicological assays were useful to screen test samples from the environment composed of various pollutants.

Keywords: cytotoxicity, phototoxicity, environmental pollutants, ecotoxicology

Introduction

Detection of environmental pollutants is very important in ecotoxicology. Environmental samples may be in the form of gas, liquid, sludge, particle and so on (Al-Khodairy and Hannan, 1997; Monarca et al., 1998; Minissi et al., 1998; Vahl et al., 1997), and this makes it difficult to examine the samples and to compare results.

Until now, mainly chemical analyses have been used to assess pollutants in water and air. Although chemical analyses are becoming more accurate due to advanced analytical technology, they detect only the targeted chemicals in the test sample, which may contain a combination of various chemicals.

Fish, birds, invertebrate animals, and plants, etc., have been used to assess the toxicity of chemicals in the environment (Huang et al.,1996; Ueda et al., 2005; Ralph and Petras, 1997; Yoshimura and Endoh, 2005; Wang et al., 2005). However, these kinds of tests aimed at assessing one target chemical in the environment and not to assess the complex mixture that exists in the environment. In view of the 3Rs, animal testing should not be conducted, if at all, without obtaining chemical information and results of *in vitro* testing first.

Meanwhile, many *in vitro* methods have been used as prescreening methods before animal testing or as alternative testing methods. We applied mammalian cell culture systems using several *in vitro* methods to detect toxicity with high sensitivity and compared the results with the chemical analyses of the samples. Since it was difficult to use all sample forms in one or a few methods, we needed to adjust the sample forms to testing method, e.g. to extract with solvent.

Two types of samples, liquid and solid, were collected from the environment. Liquid samples were used after filtration and extracts were made of the solid samples. Chemical analysis was performed on all samples as well.

Materials and methods

Liquid samples

Water samples from cooling towers on buildings and river water samples were collected. A cooling tower is a system often set on top of buildings to cool down water that has been warmed when circulating in the building. There are two types of systems, open systems and closed systems. In open systems, the warm water runs downward while air is blown through the water to cool it. It is not covered firmly, since the water is cooled down with outside air, so tiny droplets of water can spread to the environment. We collected the water of cooling towers from six facilities, four located in Tokyo, Japan and two in Kanagawa prefecture near Tokyo.

River water samples were collected from Sakawa river and its tributaries in Kanagawa prefecture (Wakuri et al., 2002). Thirteen points were chosen and water samples were collected six times over two years. At one assessment time, samples were collected from only nine points.

Colony formation test and chemical analysis were performed with these samples. The liquid samples were sterilised by filtration before use in the *in vitro* test.

Solid samples

Three types of samples, river sediments, ashes from incineration facilities, and airborne particulates were collected.

River sediments were collected in the same rivers from which the water samples were collected. We chose four points out of the points used for water collection, so that we could compare the results with those of the water analysis. The sediments were dried and extracted with organic solvent. The extracts were tested in colony formation and phototoxicity tests.

Ashes were collected from six incineration facilities. Two types of ashes, bottom ashes and fly ashes, which were collected

with an electrostatic precipitator, were obtained from four facilities. From the other two facilities, we collected bottom ashes or fly ashes. They were extracted with organic solvent and the extracts were employed in the phototoxicity test.

Airborne particulates were collected in Sapporo city, Hokkaido, which is the northernmost prefecture in Japan (Matsumoto et al., 1998). The airborne particulates that had been collected over 24 years using a high-volume air sampler and filters had been stored at -20˚C. They were extracted with organic solvent and each test sample was made up with 2 years' worth of extract. Phototoxicity test and mouse lymphoma assay were performed with these extracts.

Colony formation test

The colony formation test was used as a cytotoxicity test. BALB/3T3 A31-1-1 (BALB/3T3) cells derived from mouse embryo and VERO cells derived from the kidney of an African green monkey were used in the colony formation assay. Both cell lines, especially BALB/3T3 cells, are known to be sensitive to chemicals.

One day prior to treatment, 100 cells were seeded onto a 60-mm dish. The cells were treated with the extracts or the medium prepared with filter-sterilised water samples. The cells were treated for 7 days in case of BALB/3T3 cells and for 10 days in case of VERO cells. After incubation, cells were fixed, stained and colonies were counted.

Phototoxicity test

The *In Vitro* 3T3 NRU phototoxicity test (3T3 NRU PT), which was adopted as an OECD guideline in 2004, was used to detect phototoxicity. BALB/3T3 cells were plated into 96-well plates and the next day cells were treated with extracts.

For treatment, the medium was changed to test solution, then cells were pretreated in an incubator for 1 h. After pretreatment, one plate was irradiated at 1.7 mW UVA/cm^2 for 50 min, and another plate was kept in the dark for 50 min to assess cytotoxicity without UV irradiation.

After treatment, the cells were incubated for one more day and cell survival rates were measured by the Neutral Red uptake method.

Mouse lymphoma assay

We used a mutation assay based on L5178Y/tk$^{+/-}$-3.7.2C mouse lymphoma cells with trifluorothymidine resistance known as the mouse lymphoma assay (MLA).

The extracts were treated with or without S9 mix for three hours, then washed out. The cells were cultured for two days and then were replated to detect mutation induction.

Chemical analysis of water

A series of water analysis was performed under the Japanese water quality guideline.

We used three parameters, chemical oxygen demand (COD), iron and lead, for water from cooling towers and four parameters, number of bacteria, COD, nitrate and nitrite nitrogens, and electric conductivity for river water.

Chemical analysis of extract

The amount of polycyclic aromatic hydrocarbons (PAHs) was measured in the extracts from ashes and airborne particulates. Seven PAHs (phenanthrene, anthracene, fluoranthene, pyrene, benz(a)anthracene, benz(a)pyrene, acenaphthylene) were analysed in the ashes and eight PAHs (benz[a]anthracene, benzo[b]fluoranthene, benzo[k]fluoranthene, benzo[a]pyrene, benzo[ghi]perylene chrysene, fluoranthene, pyrene) in the airborne particulates.

Results

Water from cooling towers

The IC_{50} values in the colony formation tests with BALB/3T3 and VERO cells were comparable in the water samples of four of the six buildings. The other two samples showed no toxicity in one cell line and strong toxicity in the other. This suggested that the two cell lines display differing sensitivities to some substances.

The results of the chemical analysis of lead, iron and COD showed tendencies similar to the colony formation tests. The samples that showed higher contents of lead, iron or COD, which means that the waters were "dirty", also showed stronger cytotoxicity.

Water from rivers

After obtaining all results, the water samples were ranked. First the ranking was made for each time point for each parameter. The ranking numbers of two consecutive years were summed-up for each point, then they were renumbered according to their sum. Ranks were also determined for overall cytotoxicity and for overall chemical content.

Though the regression coefficient between the results for BALB/3T3 and VERO cells was less than 0.6, and the correlation between the cytotoxicity ranking and each single chemical parameter was about 0.5-0.8, the total ranking of cytotoxicity and the overall chemical content showed a good correlation with a regression coefficient of 0.9.

Sediments from rivers

VERO cells showed higher sensitivity (2-5 times as IC_{50} values) than BALB/3T3 cells for all four sample points. However, the sediments collected at upper stream sample points showed less cytotoxicity than those collected at lower points in the stream in both cell lines. This tendency was also reflected in the results of the chemical analysis of the water collected at the same points.

Meanwhile, the results of the phototoxicity assay were similar to those of the cytotoxicity test in the upper three points. The lowest point showed less phototoxicity than the middle two points. This might be because the sediment material was different from the middle two points. The sediments at the middle two points were mud-like but the lowest point included more sand-like particles. This implies that the absorption conditions could have been different and, therefore, the results did not reflect the water analysis.

Ashes from incineration facilities

Ten ash samples from the incineration facilities were examined. Fly ashes contained more PAHs than bottom ashes except for those of one facility. In some facilities, the extracts showed stronger phototoxicity, as ashes contained large amount of PAHs. However, the extract that contained the highest amount of PAHs was the least cytotoxic under irradiation conditions and almost the least cytotoxic under non-irradiation conditions.

Airborne particulates

In the phototoxicity test, cytotoxicity without irradiation was stable in the samples collected over 24 years. On the other hand, cytotoxicity under irradiation decreased over the 24 years in the same manner as the amount of PAHs decreased.

Mutation induction in the presence of S9 mix showed the same downward trend as the amount of PAHs and cytotoxicity under irradiation. There were no clear tendencies regarding mutation induction in the absence of S9 mix, though the old samples (first 6-8 years) showed higher mutation induction.

Discussion

The cytotoxicity of liquid samples from cooling towers and rivers displayed the same tendency as the chemical analysis. The two cell lines used did not show a high correlation, but the difference in their sensitivity can be thought of as a differing ability to detect some pollutants. Since the chemical analysis of the water has been used to judge the water quality, *in vitro* cytotoxicity testing could be used to detect water pollution, if the cell line is chosen carefully.

Regarding the solid samples, the cytotoxicity of river sediments corresponded with the water analysis, though phototoxic effects showed some discrepancies. One reason might be that we did not analyse the extracts themselves and phototoxicants could be accumulated in sediments. Another is that the water analysis did not include detection of contaminants that are phototoxic.

Phototoxicity results of ashes did not agree well with the results of the chemical analyses either. Though PAHs are known as strong phototoxicants (Yu, 2002; Nikolaou et al., 1984), the PAHs we analysed might not be the major contributors to the phototoxicity or the major substances causing phototoxicity might not have been analysed. Since the conditions of incineration and the rubbish burnt were different in each facility, the contents of the ashes could be different as well.

Airborne particulates were tested by phototoxicity and mutation assay. Both results suggested that PAHs contributed part of the activities. PAHs are well-known to show photo-enhanced cytotoxicity and mutation reaction with S9 mix (Penning et al., 1999). In our results, cytotoxicity under irradiation conditions, mutation induction in the presence of S9 mix and the amount of PAHs displayed the same tendencies. This suggests that PAHs were the main cause of phototoxicity and mutation induction in airborne particulates.

Conclusion

Firstly, the results of the *in vitro* testing generally agreed with the results of the water/chemical analysis. Differences might have been caused if the substances analysed chemically did not induce toxic effects *in vitro*.

Secondly, *in vitro* tests can be used to detect various biological reactions by testing a complex sample mixture with several test methods. The chemical analysis shows only the quantity of target substances and not the effects of the substances, whereas the *in vitro* tests show biological reactions that we could use to predict the *in vivo* reaction.

Finally, we expect that *in vitro* testing would make it easier to compare the results from different environmental samples without requiring any animal testing.

References

Al-Khodairy, F. and Hannan, M. A. (1997). Exposure of organic extracts of air particulates to sunlight leads to metabolic activation independence for mutagenicity. *Mutat. Res. 391*, 71-77.

Huang, X., Zeiler, L. F., Dixon, D. G. and Greenberg, B. M. (1996). Photoinduced toxicity of PAHs to the foliar regions of Brassica napus (Canola) and Cucumbis sativus (Cucumber) in simulated solar radiation. *Ecotoxicol. Environ. Saf. 35*, 190-197.

Matsumoto, Y., Sakai, S,. Kato, T. et al. (1998). Long Term Trends of Particulate Mutagenic Activity in the Atmosphere of Sapporo. 1. Determination of Mutagenic Activity by the Conventional Tester Strains TA98 and TA100 during an 18-Year Period (1974-1992). *Environ. Sci. Technol 32*, 2665-2671.

Minissi, S., Caccese, D., Passafiume, F. et al. (1998). Mutagenicity (micronucleus test in Vicia faba root tips), polycyclic aromatic hydrocarbons and heavy metal content of sediments collected in Tiber river and its tributaries within the urban area of Rome. *Mutat. Res. 420*, 77-84.

Monarca, S., Zanardini, A., Feretti, D. et al. (1998). Mutagenicity and clastgenicity of gas stove emissions in bacterial and plant tests. *Environ. Mol. Mutagen. 31*, 402-408.

Nikolaou, K., Masclet, P. and Mouvier, G. (1984). Sources and chemical reactivity of polynuclear aromatic hydrocarbons in the atmosphere – A critical review. *The Science of the Total Environment 32*, 103-132.

Penning, T. M., Burczynski, M. E., Hung, C. et al. (1999). Dihydrodiol dehydrogenases and polycyclic aromatic hydrocarbon activation: generation of reactive and redox active o-quinones. *Chemical Research in Toxicology 12*, 1-18.

Ralph, S. and Petras, M. (1997). Genotxicity monitoring of small bodies of water using two species of tadpoles and the alkaline single cell gel (Comet) assay. *Environ. Mol. Mutagen. 29*, 418-430.

Ueda, H., Ebino, K., Fujie, H. et al. (2005). Reassessment of ecotoxicities of 1,1,1-trichloro-2,2-bis(4-chlorophenyl)ethane

in Japanese quail under the OECD draft new avian one-generation reproduction test guideline. *Regul. Toxicol. Pharmacol. 41*, 167-174.

Vahl, H. H., Karbe, L. and Westendorf, J. (1997). Genotoxicity assessment of suspended particulate matter in the Elbe river: comparison of Salmonella microsome test, arabinose resistance test, and umu-test. *Mutat. Res. 394*, 81-93.

Wakuri, S., Takahashi, A., Hayashi, M. and Tanaka, N. (2002). Cytotoxicity test with waters from Sakawa rivers and its tributaries. *Journal of Japan Society on Water Environment 25*, 669-674.

Wang, H. Y., Olmstead, A. W., Li H. and Leblanc, G. A. (2005). The screening of chemicals for juvenoid-related endocrine activity using the water flea Daphnia magna. *Aquat Toxicol. 74*, 193-204.

Yoshimura, H. and Endoh, Y. S. (2005). Acute toxicity to freshwater organisms of antiparasitic drugs for veterinary use. *Environ. Toxicol. 20*, 60-66.

Yu, H. (2002). Environmental carcinogenic polycyclic aromatic hydrocarbons: photochemistry and phototoxicity. *Journal of Environmental Science and Health Part C – Environmental Carcinogenesis & Ecotoxicology Review C20*, 149-183.

Correspondence to
Shinobu Wakuri
Hatano Research Institute
Food and Drug Safety Center
729-5 Ochiai
Hadano, Kanagawa 257-8523, Japan
e-mail: wakuri.s@fdsc.or.jp

Workshop 5.11
Mechanisms of chemically-induced ocular injury and recovery

Ocular Toxicology *In Vitro* – Cell Based Assays

Monica Berry and Marcus Radburn-Smith
Academic Unit of Ophthalmology, University of Bristol, UK

Summary
Interactions between the three cell types in the cornea control differentiation and responses to stimuli. We have sequentially added cell types in a 3-dimensional construct to assess the minimal requirements for a toxicology model of the human cornea. Co-cultured cell types displayed patterns of cytokines different from the single cell-type 3D models. Following exposure to toxicants there were marked changes in cytokine profiles, related to the toxicant and markedly influenced by the epithelial cell-line used. For a rational choice of cell lines, their steady-state patterns of signal molecules should be compared to those in normal human preocular fluid.

Keywords: cornea, cytokines, human, in vitro, toxicology

Introduction and questions

The health of the ocular surface is intimately linked with quality of vision, and is also a determining factor in the success of surgical procedures. The ocular surface encompasses the outer structures of the eye (cornea, conjunctiva, and lids), the lacrimal gland and tear film and the innervation and immune structures that link them (Stern et al., 2004). Interactions within and between these components maintain the barrier to the environment and modulate responses to external stimuli (Cook et al., 2001; Stramer et al., 2003; Wilson et al., 2003; Hazlett, 2004; Holan et al., 2004; Wilson et al., 2004; Lema and Duran, 2005; Narayanan et al., 2005).

The Colipa Strategy for the Development of *in vitro* Alternative Methods focuses on physiological processes – from organ level to gene activation – that can be used to monitor responses to toxicants and recovery from injury. Among the programme aims, central to the replacement R of the 3Rs and to cell based assays, is identifying quantitative endpoints predictive of the nature and severity of injury. The ultimate aim is to replace the Draize test – currently the gold standard of ocular toxicity testing – with a scientifically valid animal-free strategy.

We needed to establish the minimal degree of complexity which is necessary and sufficient to represent the human cornea in cell culture models of toxicity. We have assessed the behaviour of undisturbed cultures and the response to a single toxicant application in constructs from monolayers to three-dimensional cultures comprising stratified human corneal epithelia and quiescent human stromal cells, to probe stability and correlates to injury and recovery.

Starting from the premise that depth (cell types) and area are the defining characteristics of an injury (Jester et al., 1998a; Jester et al., 1998b; Jester et al., 2000; Jester et al., 2001), we have explored endpoints related to metabolism and barrier function and sought correlations with patterns of cytokines, mediators that are known to be secreted by, and influence, the different cell types of the ocular surface. For the refinement R of the 3Rs and, more importantly, to anchor the *in vitro* tests with *in vivo* physiology, the add-on benefit of the latter endpoints is that signal molecules are readily detectable in (human) tears, and available through non-invasive tests.

Materials and methods

Epithelium

In addition to availability and stability with passage number, minimum *a priori* criteria were set for choosing human corneal

or conjunctival epithelial cell lines: ability to stratify and establish a barrier to penetration. Acute toxicity is correlated with barrier function (Konsoula and Barile, 2005). The corneal stroma is protected by the tight barrier of epithelium, which is in turn protected by the preocular fluid. A number of corneal and conjunctival epithelial cell lines were assessed for stratification, morphological differentiation and trans-epithelial electrical resistance (TEER).

Initially, epithelia were cultured according to originator's instructions: the cultures were then transferred and maintained in fully defined medium without antibiotics (KGM Bulletkit, Cambrex Biosciences, NJ, USA). Stratification was achieved by supplementing the medium with 1.25 mM $CaCl_2$ and cultured at liquid-air interface. The choice of serum-free medium follows from the need for minimally/non-activated stromal cells. This medium will not interfere with toxic effects and endpoint measurements.

Stroma

With approval of the Committee for Ethics in Research and consent for use in research, donor corneas unsuitable for transplantation were used to generate primary human corneal fibroblasts. These were amplified in DMEM with 10% foetal bovine serum (Invitrogen, Paisley, UK) and cryopreserved. Cells used to populate a collagen type I gel were maintained for at least 1 week in fully defined medium to inhibit cell proliferation and activated phenotype. Before seeding in collagen gels the cells were dissociated using trypsin. Soybean trypsin inhibitor (Sigma, Poole, UK) was used to stop the enzyme.

The matrix was built using 3 mg/ml bovine skin type I collagen (Vitrogen, Angiotech Biomaterials, Ca., USA), gelling at pH 7.4 (Taliana et al., 2000). Cells were included in the collagen solution at a concentration of 6 x 10^4 cells/ml. The gels, volume 0.25 ml per 24 well insert, were cast at their final placement and maintained in fully defined medium for 2 weeks before any further manipulation.

Epithelial-stromal construct

At least 2 weeks after matrix formation, epithelial cells were seeded at 10^5 cells/gel. The construct was maintained for 2 days and then lifted to liquid-air interface, where it was cultured for a further 7 days.

Toxicants and exposures

The choice of toxicants unifies the Colipa project: they are all from the same batch and are distributed to the collaborating laboratories by Colipa. We tested NaOH (Riedel-de-Haën, Germany), Sodium dodecyl sulphate (SLS, OmniPur®, Merck, Germany), and the non-ionic surfactant Tomadol 45-7 (Tomah, Ca. USA), each prepared in tissue culture water (Sigma). The concentration chosen was 0.66%, at which NaOH produces an initial lesion covering approximately three quarters of the surface of a stratified epithelium. For monolayers, this concentration was halved to take into account the smaller mass of the construct.

All constructs were exposed to 10 µl toxicant applied topically for 10 minutes and followed by an exhaustive wash with 1.0 ml tissue culture medium. Cultures were followed at 1 and 4 h for early signs of injury and then at 1, 3, 5, and 7 days post-exposure to evaluate the ensuing dynamics of responses to injury and recovery.

Assays

Trans-epithelial electrical resistance: TEER was measured with Millicel Electrodes (Millipore, Ma., USA) held in a fixed support. A cell-free insert was included as reference at each measurement.

Protein: After solubilisation with 1% Triton X-100 (Sigma) overnight at 4ºC, protein was quantified using the bicinchoninic acid kit (BCA™ Protein Assay Kit, Pierce, IL, USA) according to manufacturer's instructions.

Alamar Blue: This reagent (AlamarBlue, Serotec, Oxford, UK) was used at 2.5%, a concentration with minimal toxicity on repeat cell loading, every other day for 14 days. Fluorescence was measured using a Spectramax fluorimeter (Molecular Devices Corporation, Ca., USA) at $\lambda_{excitation}$ = 545 nm and $\lambda_{emission}$ = 590 nm, and expressed in arbitrary units. The sensitivity of the instrument was fixed for all readings.

Cytokines: TNF, IL-1β, IL-6, IL-8, IL-10 and IL-12p70 were measured simultaneously in each supernatant sample using the BD Cytometric Bead Array (BD Biosciences, Ca., USA) on a FACSCalibur (BD Biosciences). The fluorescence of samples and standard dilutions was assessed at least twice in each assay. Results are presented as means of a number of cultures, while the mean of replicate readings counted as 1. The dynamic range of the assays were 5-5000 pg/ml for IL-6 and IL-8.

Results

Some cell lines did not fulfil the minimal criteria for inclusion in tests and were not investigated further. For example, the immortalised human corneal epithelial cell line (Canadian cells, gift from May Griffith) reached only a fraction of the TEER achieved by the immortalised human corneal epithelial cells (USA cells, donated by Ilene Gipson), as shown in figure 1. The TEER of IOBA-normal human conjunctival cells (gift from Yolanda Diebold) was equal to that of the Canadian cells. Note that *in vivo* the conjunctiva is much leakier than the cornea.

It is interesting to note that the cytokine levels were specific to each cell line: in Araki-Sasaki supernatants IL-6 and IL-8 were in the order of 30 pg/ml and 400 pg/ml respectively, while in the Canadian cells levels were around 1500 and 2000 pg/ml respectively. In contrast, the USA cells secreted no detectable IL-6 when stratified. There was a clear time dependence of cytokine secretion, strongly influenced by the culture model. Stratification and culture at liquid-air interface, each affected the pattern and secretion of cytokines.

Further changes were observed when epithelia were cultured with keratocyte-populated stroma. An advantage of using a collagen matrix that can be altered by the cells growing therein becomes evident in that shrunken collagen gels can be easily identified. The difference in cytokine secretion from shrunken and unaltered gels is an indication of interactions between stro-

mal and epithelial cells (fig. 2), and stromal and endothelial cells (cell line gift from May Griffith). IL-8 concentrations also changed in response to the above factors. No cytokines were detected in supernatants of keratocyte gels.

When epithelial and endothelial cells were co-cultured with medium separating the two cell types, we observed a larger influence on IL-6 and IL-8 cytokine secretion than when in direct contact. Concentrations and magnitude of change with two cell types also depended on the medium used.

Preliminary experiments suggested that the optimum experimentation period would commence seven days after air-lifting and that the stratified epithelium would show no signs of attrition for seven days thereafter. We compared the different constructs for the projected duration of the experiment by measuring a metabolic correlate with Alamar Blue, protein as a measure of cell proliferation, and cytokine production. For monolayer cultures the starting point was taken at confluence, judged by microscopic inspection. While there was a continuous increase in cell mass (protein concentration) in the monolayer, the absolute redox activity remained stable, suggesting a decrease per cell. Protein concentrations and reductive activity were decreased in the USA stratified epithelia after day 5 of the experiment, while the epithelial-stromal constructs were stable for 7 days. Cytokine production per unit mass increased in monolayers, while in the stratified epithelia and the epithelial-stromal constructs cytokine levels were much lower and little changed during the experiment (fig. 3). Remarkably, IL-6 production was not quantifiable in stratified epithelia, and the levels of IL-8 were also below those of constructs containing a stromal cell seeded gel and stratified epithelia.

In response to single, short-term toxicant applications all constructs showed decreases in total reductive activity and cell mass (protein concentration). However, the time course and extent of these changes varied according to the construct: the non-ionic surfactant caused a very early and devastating decrease in cell number and activity in the monolayer. These losses of cells and metabolic capability occurred later in stratified epithelia; they were smaller and preceded by a short-term increase in reductive activity in epithelial-stroma constructs. In the case of this toxicant, an increasing secretion of IL-6 and IL-8 was measured towards the end of the experiment in the epithelial-stromal construct, perhaps suggestive of some recovery.

Discussion and conclusions

In response to public demand and European legislation, testing of cosmetic ingredients on live animals will cease in the near future. A consensus has emerged that any methodology that is developed to replace the use of animals in ocular toxicity testing has to fulfil a number of *a priori* criteria in order to be considered for pre-validation. These criteria include scientific purpose, mechanistic basis, statement of limitations and appropriate controls.

Within the Colipa Strategy for Development of *in vitro* Alternative Methods programme we have aimed at paradigms that address these criteria. Damage to the cornea is the most serious in functional visual terms, hence it was chosen as the target organ representing the ocular surface. We have reasoned that mediators of intercellular communication that can also be involved in inflammation would be a mechanistically-linked correlate of the response of the cornea to external stimuli. Inflammation is an aggravating result of ocular injury, and an integrating response of the ocular surface.

The results presented here indicate that cytokine production

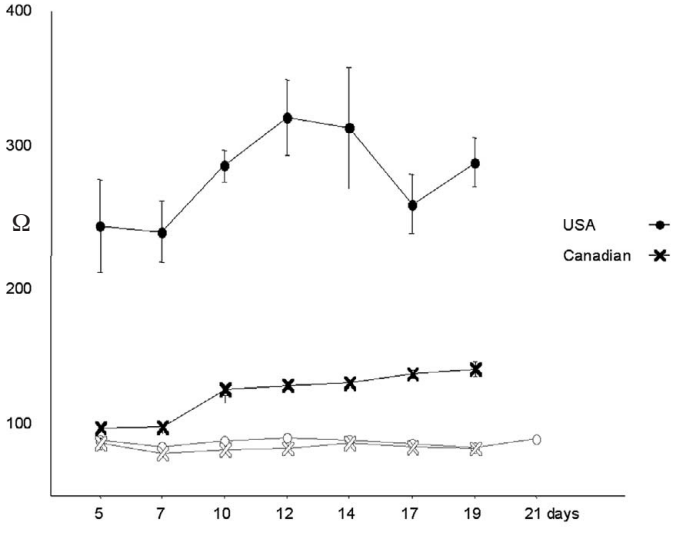


Fig. 1: Trans-epithelial electrical resistance
USA cells: filled circles; Canadian cells: filled crosses; blank inserts: grey, unfilled symbols

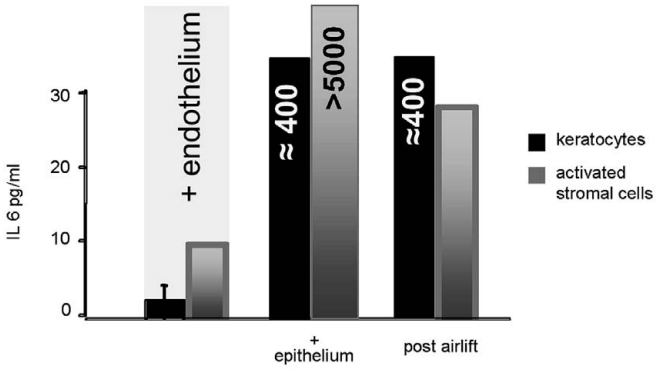


Fig. 2: IL-6 levels in different culture conditions
Black bars: quiescent keratocytes; grey bars: activated keratocytes (shrunken gel)

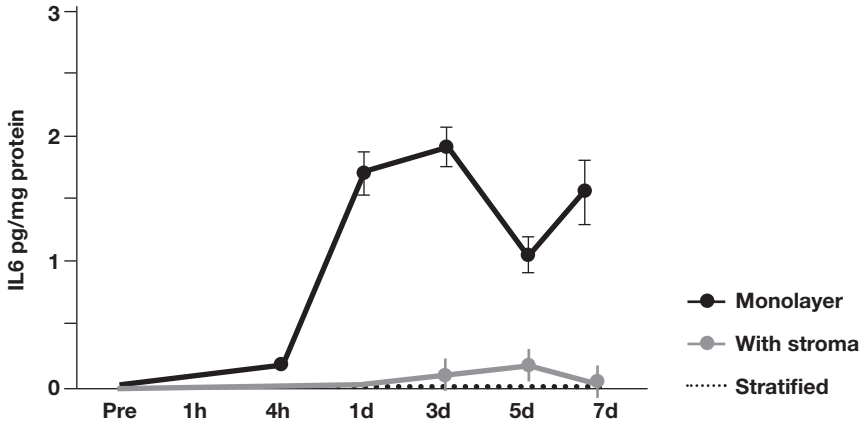


Fig. 3: IL-6 secretion in cell culture constructs with USA epithelia and quiescent keratocytes
Black line: monolayer; grey line: epithelialised stroma; dashed line: stratified epithelium

reflects the type and three-dimensional arrangement of cultured cells. Furthermore, the stability of the constructs is enhanced by the presence of different cell types. This indicates that a stratified epithelium alone is not a sufficient model for corneal toxicity tests. Elegant three-dimensional human corneal epithelial constructs, e.g. Skinethic (Van Goethem et al., 2005), could be combined with a three-dimensional stroma. Responses to a small number of chemicals suggest that the epithelial-stromal model allows some recovery to be studied and might be more discriminating than the single cell type models. In this construct, assessment of the time course of cytokine secretion after injury pointed to changes in cell activity that were not detectable by measuring reductive activity or cell mass.

There remains a caveat to the choice of cell lines in corneal constructs: each of the cell lines tested secreted different levels of cytokines. A comparison with data on cytokines in human tears is complicated by the fact that such data have not been obtained under standardised conditions, even when a healthy control group was included in the study. Our results indicate that after a chemical injury IL-6 increases in tears, compared to unexposed unrelated eyes (Berry and Jeffreys, 2001). In animals, IL-6 and IL-8 (and IL-1α) were elevated for a period of days after exposure to a toxicant, each cytokine taking a distinctive time course (Den et al., 2004). *In vitro*, twelve borderline irritation eye make-up removers increased IL-8 levels in supernatants of centrifuged CEPI monolayers from 1 to 23 times over control values (Debbasch et al., 2005), suggesting increased sensitivity of this endpoint to toxicant effects.

After sequentially building more complex corneal models and comparing their behaviour in steady state and after exposure to toxicants, we conclude that a three dimensional culture involving the three major cell types in the cornea is a stable construct and potential model for toxicity testing. Concentrations of toxicants used in these model-development studies have been at least one or two orders of magnitude smaller than those used in ex vivo studies done by our collaborators. This encourages us to further focus on mild to moderate toxicants when exploring the potential and limitations of this methodology.

References

Berry, M. and Jeffreys, D. (2001). Ocular injuries from household chemicals: Early signs as predictors of recovery. *In vitro and molecular toxicology 14*, 5-13.

Cook, E. B., Stahl, J. L., Lowe, L. et al. (2001). Simultaneous measurement of six cytokines in a single sample of human tears using microparticle-based flow cytometry: allergics vs. non-allergics. *J. Immunol. Methods 254*, 109-118.

Debbasch, C., Ebenhahn, C., Dami, N. et al. (2005). Eye irritation of low-irritant cosmetic formulations: correlation of in vitro results with clinical data and product composition. *Food Chem. Toxicol. 43*, 155-165.

Den, S., Sotozono, C., Kinoshita, S. and Ikeda, T. (2004). Efficacy of early systemic betamethasone or cyclosporin A after corneal alkali injury via inflammatory cytokine reduction. *Acta Ophthalmol. Scand. 82*, 195-199.

Hazlett, L. D. (2004). Corneal response to Pseudomonas aeruginosa infection. *Progress in Retinal and Eye Research 23*, 1-30.

Holan, V., Vitova, A., Pindjakova, J. et al. (2004). Corneal stromal cells selectively inhibit production of anti-inflammatory cytokines by activated T cells. *Clin. Exper. Immunol. 136*, 200-206.

Jester, J. V., Li, H. F., Petroll, W. M. H. et al. (1998a). Area and depth of surfactant-induced corneal injury correlates with cell death. *Invest. Ophthalmol. Vis. Sci. 39*, 922-936.

Jester, J. V., Li, L., Molai, A. and Maurer, J. K. (2001). Extent of initial corneal injury as a basis for alternative eye irritation tests. *Toxicol. In Vitro 15*, 115-130.

Jester, J. V., Molai, A., Petroll, W. M. et al. (2000). Quantitative characterization of acid- and alkali-induced corneal injury in the low-volume eye test. *Toxicol. Pathol. 28*, 668-678.

Jester, J. V., Petroll, W. M., Bean, J. et al. (1998b). Area and depth of surfactant-induced corneal injury predicts extent of subsequent ocular responses. *Invest. Ophthalmol. Vis. Sci. 39*, 2610-2625.

Konsoula, R. and Barile, F. A. (2005). Correlation of in vitro cytotoxicity with paracellular permeability in Caco-2 cells. *Toxicol. In Vitro 19*, 675-684.

Lema, I. and Duran, J. A. (2005). Inflammatory molecules in the tears of patients with keratoconus. *Ophthalmology 112*, 654-659.

Narayanan, S., Glasser, A., Hu, Y. S. and McDermott, A. M. (2005). The effect of interleukin-1 on cytokine gene expression by human corneal epithelial cells. *Exp. Eye Res. 80*, 175-183.

Stern, M. E., Gao, J., Siemasko, K. F. et al. (2004). The role of the lacrimal functional unit in the pathophysiology of dry eye. *Exp. Eye Res. 78*, 409-416.

Stramer, B. M., Zieske, J. D., Jung, J. C. et al. (2003). Molecular mechanisms controlling the fibrotic repair phenotype in cornea: Implications for surgical outcomes. *Invest. Ophthalmol. Vis. Sci. 44*, 4237-4246.

Taliana, L., Margaret, Evans, M. D. M., Slobodan, D. et al. (2000). Vitronectin or fibronectin is required for corneal fibroblast-seeded collagen gel contraction. *Invest. Ophthalmol. Vis. Sci. 41*, 103-109.

Van Goethem, Adriaens, E., Alepee, N. et al. (2005). Prevalidation of a new in vitro reconstituted human cornea model to assess the eye irritating potential of chemicals. *Toxicol. In Vitro*, in press.

Wilson, S. E., Mohan, R. R., Netto, M. et al. (2004). RANK, RANKL, OPG, and M-CSF expression in stromal cells during corneal wound healing. *Invest. Ophthalmol. Vis. Sci. 45*, 2201-2211.

Wilson, S. E., Netto, M. and Ambrosio, R. Jr. (2003). Corneal cells: Chatty in development, homeostasis, wound healing, and disease. *Am. J. of Ophthalmol. 136*, 530-536.

Acknowledgements

This project is part of the Colipa Strategy for the Development of *in vitro* Alternative Methods. We thank Colipa for financial support and the Eye Irritation Task Force members for the choice and procurement of toxicants and for many stimulating discussions throughout the project. We would like to thank Leon Bruner (Gillette), Richard Clothier (University of Nottingham), Yolanda Diebold and Margarita Calonge (IOBA, Valladolid), Dan Dimitrijevitch (University of North Texas), Ilene Gipson and Sandy Spurr-Michaud (Schepens Institute, Boston) and May Griffith (University of Ottawa) for gifts of cell lines and methodological advice. This project is part of the research undertaken by Marcus Radburn-Smith for a PhD degree at the University of Bristol. We are grateful to the National Eye Research Centre for additional support for this project.

Correspondence to

Dr Monica Berry
University of Bristol
Academic Unit of Ophthalmology
Bristol Eye Hospital
Lower Maudlin St
Bristol BS1 2LX
UK
e-mail: mon.berry@bristol.ac.uk

Can Toxicogenomics be Used to Identify Chemicals that Cause Ocular Injury?

Mike Boulton[1,2] and Mike Wride[2]

[1]Department of Ophthalmology & Visual Sciences, The University of Texas Medical Branch, Galveston, Texas, USA;
[2]School of Optometry and Vision Sciences, Cardiff University, Cardiff, UK

Summary

With the impending ban by the European Union on the use of the Draize eye test for cosmetic testing and to meet concerns regarding the 3Rs, an in vitro replacement is essential. This article considers a toxicogenomic approach that will allow a gene fingerprint to be produced from chemical exposure of corneal constructs prepared from human cell lines. Differential fingerprint profiles will allow identification of mild, intermediate and severe toxic preparations. This toxicogenomic approach will provide a rapid, high throughput, accurate diagnostic assay for the effects of toxic chemicals on the eye.

Keywords: Draize eye test, toxicogenomics, cornea

Introduction

The Draize eye irritation test has been mandated for routine use by manufacturers to study the safety of their products for marketing. This test, which is currently the gold standard, has several limitations:

a) It lacks objective and accurate quantification
b) It requires large numbers of animals
c) It is a slow, expensive, and a time consuming process
d) It is opposed by animal welfare groups as well as increasing numbers of the general public concerned about "non-medical" animal experimentation.

A total of 3,300 rabbits were used for eye irritation studies in the UK in 1994, of which the majority were used for cosmetics testing. It was recommended that the use of the Draize eye irritation test should be phased out within Europe by 1997, providing that suitable *in vitro* alternatives had been developed. Despite this not being the case, the European Parliament has subsequently decreed that the Draize eye test will no longer be allowed from 2009. This has left the manufacturing and cosmetic companies with a dilemma, since they must ensure safety of all new products entering the market. In an attempt to overcome these problems, alternative *in vitro* toxicity tests have been sought. Although a number of *in vitro* methods have been proposed as alternatives to the *in vivo* Draize test, they are either overcomplicated or oversimplified to such an extent that they measure only a single aspect of eye irritation (e.g. cell death) (Prinsen, 2005). Such tests are not considered sufficiently rigorous or informative, especially for moderate or mildly toxic chemicals. One possible alternative is to use a toxicogenomic approach to provide a gene fingerprint directory (i.e. pattern recognition), which will identify the toxicity of chemicals in corneal constructs. If successful, this will allow the production of specific diagnostic gene arrays for high throughput assays of all potentially toxic preparations that enter the market.

Chemical injuries to the eye

The first points of contact in chemical exposure to the eye are the eyelids, conjunctiva and cornea. Severe injury can result in corneal scarring and neovascularisation, which leads to opacification of the cornea, thus preventing light reaching the retina (Khaw et al., 2004). In extreme cases, treatment may require corneal transplantation, keratoprosthesis or limbal cell transplantation (Dogru and Tsubota, 2005). If these avenues are unsuccessful, the patient will suffer blindness or severe visual impairment. Mild chemical injury usually results in irritation and reddening. This either recovers by itself or requires some antibiotic and steroid treatment. Moderate chemicals initiate intermediate tissue damage and the outcome is dependent on the nature of the chemical (e.g. alkali, acid, detergent) and duration of exposure.

Not surprisingly, the nature of tissue damage and biological responses is very different between severe and mild chemical injury. Severe chemical injury to the cornea will result in tissue coagulation and large-scale cell death via necrosis. Immediate effects will be on the epithelium, but penetration into the stroma will cause damage to Bowman's membrane and the underlying stroma with loss of keratocytes and disorganisation of the regular collagen lamellae. This results in breakdown of barrier function, making the cornea more prone to infection, and destabilisation of corneal hydration leading to swelling. The release of cytokines and other mediators will attract macrophages, leading to an acute inflammatory response and neovascularisation. The immediate survival response is rapid proliferation of cells and disorganised matrix deposition, which often results in corneal scarring. Ulceration and corneal liquefaction can result, due to sustained upregulation of matrix metalloproteinases (MMPs). Not surprisingly, the biological responses associated with severe corneal damage result in a massive temporal change in gene expression. Furthermore, tissue remodelling can continue for up to 3 years post injury.

By contrast, mild chemical injury to the eye will result in limited tissue damage and a normal repair response. There is usually some cell loss, normally of the corneal epithelium, which is usually repaired within 72 hours through the division of resident transit amplifying cells and limbal stem cells. There will be some pain and irritation due to damage to the nerve endings in the epithelial layer. The mild associated inflammatory response will result in a reddening of the eye, which may need treatment with steroids and antibiotics. As with severe chemical injury, there will be a temporal change in gene expression, but the gene expression profile will be very different to that for severe injury.

While mild chemical injury does not normally impair vision, the discomfort to the individual is unacceptable and should be avoided if possible. Thus, it is important that manufacturers, employers and Health and Safety Executives are fully aware of potential chemical contraindications, can advise on protection and can recommend treatment following accidental exposure.

Feasible alternatives to the Draize test

Numerous alternatives to the Draize test have been proposed. These include cell culture, *ex vivo* organ culture and corneal reconstructs. While cell cultures (corneal or non-ocular, such as skin) can provide valuable information on severely toxic chemicals (the cells simply die!), they are limited in that they are usually composed of a single cell monolayer and, thus, do not measure tissue penetration. Such cultures have been used to assess the release of cytokines, growth factors and other mediators as a function of inflammation, but validation for mild or moderate chemicals is difficult due to the simplicity of the test.

Ex vivo organ culture using bovine, rabbit and porcine corneas has been attempted with outcome measures including morphology, release of biological mediators and opacification. Such models have the advantages that they have the typical 3D corneal structure, consist of numerous cell types and allow chemical penetration to be taken into account. Limitations include access to sufficient numbers of corneas, quality control (there will be considerable biological variation), the lack of a rapid throughput assay and the fact that these tissues are not human.

Corneal reconstructs (either epithelium/stroma or epithelium/stroma/endothelium) can be prepared in culture from established and well-characterised human cell lines (Griffith et al., 1999). These constructs, which produce an excellent 3D corneal architecture, allow excellent quality control and permit chemical penetration to be assessed. The limitation of these models to date is the nature of the outcome to be used to monitor mild or moderate chemical injury. Outcome measurements have tended to be similar to those described for organ culture and include morphological assessment and the measurement of the release of mediators. However, an alternative and more reliable outcome would be to monitor global gene expression changes in a construct following chemical exposure.

Aim

The strategy for the global gene expression approach is to develop a gene fingerprint directory (a pattern recognition approach), which will identify chemicals and preparations toxic to a bioengineered human cornea and that will be at least as sensitive as the Draize test.

Specific aims will include:

- Determination of different genetic fingerprints of human, bioengineered corneas exposed to different generic groups of toxic chemicals using microarray analysis.
- Recommendations for a set of selected human genes (<100) involved in eye irritation that are differentially expressed in the microarray experiments and are therefore important for the production of specific "diagnostic array" chips.

The proposed genomics assay for eye irritation will focus on pattern recognition rather than individual changes in genes. This pattern recognition approach, which depends on dynamics, dose and kinetics, will identify markers, e.g. for inflammatory processes, that may be exploited in the development of other *in vitro* assays.

Experimental design

For proof of principle, bioengineered human corneal constructs will be exposed to a range of generic chemicals (e.g. alkali, acid, detergent) with varying degrees of known ocular damage (severe, moderate, mild) and penetration, at varying concentrations, for different time periods. Affymetrix Gene Array chips will then be used to investigate changes in gene expression – using the U133A and B chips containing known and unknown genes (Wilson et al., 2002). A broad-based gene approach is essential in the first instance to avoid missing key diagnostic genes. A thorough statistical analysis using standard Affymetrix array analysis software will be undertaken and genes divided into clusters based on their functional categories. Statistically significant differences in gene expression profiles are expected for particular categories of genes under each experimental condition. Thus, it will be possible to define a gene expression fingerprint for the effect of each toxic chemical in this *in vitro* system to monitor tissue damage and repair.

Based on the gene expression profiles obtained, it is proposed to produce specific chips ("diagnostic arrays") containing selected genes that are differentially expressed in our human *in vitro* model microarray experiments. Thus, a rapid high-throughput diagnostic assay using bioengineered human corneas will be developed, which will give a global fingerprint of gene expression for the effects of toxic chemicals on the cornea. A subset of human genes will be selected that are consistently differentially expressed according to the arrays under the different experimental conditions. The possibility also exists to identify genes whose expression is consistently altered in different *in vitro* models (these genes will also be confirmed as differentially expressed

using RT-PCR and/or quantitative PCR). Individual genes and/or functional categories of genes, which are differentially expressed in all experimental paradigms, will be prioritised for inclusion on a human gene "diagnostic array", which will allow for identification of "toxicity fingerprints" in diagnostic arrays. Key diagnostic genes will include markers for inflammatory processes, apoptosis, and those for regeneration, and/or embryonic development. Selected genes will be arrayed onto glass slides for diagnostic array analysis. Appropriate array analysis software will then be used to compare the gene expression profiles obtained from diagnostic tests using cosmetics with the gene expression profiles generated by known toxic chemicals.
Of particular importance will be

1. the concentration of chemical to be tested, since potency is likely to vary significantly between *in vitro* and *in vivo*.
2. duration of exposure – minutes or hours?
3. number of exposures, should there be repeat exposures and should there be irrigation as would occur with tearing *in vivo*?
4. the timing of analysis. Since toxicity will involve both damage and repair, there will be a temporal change in the gene expression profile. Thus, the most comprehensive approach will be to look at expression profiles of early and late response genes.

Validation

The toxicogenomics model will require validation at a number of levels:

1. comparison with a database on chemical injury in humans
2. comparison with toxicity results obtained with the Draize test
3. approval by the relevant bodies (e.g. ECVAM, ICCVAM, NICEATM)

Conclusion

Toxicogenomics offers the possibility to produce a rapid, high throughput, accurate diagnostic assay for the effects of toxic chemicals on the eye. The approach, once developed and validated, will offer a high level of quality control, will be highly reliable, and will be quick and simple. Furthermore, this assay will meet the requirements of the European directive to phase out the Draize test by 2009 and should significantly reduce the number of animals used in this procedure.

References

Dogru, M. and Tsubota, K. (2005). Current concepts in ocular surface reconstruction. *Semin. Ophthalmol. 20 (2)*, 75-93.

Griffith, M., Osborne, R., Munger, R. et al. (1999). Functional human corneal equivalents constructed from cell lines. *Science 286 (5447)*, 2169-2172.

Khaw, P. T., Shah, P. and Elkington, A. R. (2004). Injury to the eye. *B.M.J 328 (7430)*, 36-38.

Prinsen, M. K. (2006). The Draize Eye Test and in vitro alternatives; a left-handed marriage? *Toxicol. In Vitro 20*, 78-81.

Wilson, A. S., Hobbs, B. G., Speed, T. P. and Rakoczy, P. E. (2002). The microarray: potential applications for ophthalmic research. *Mol. Vis. 8*, 259-270.

Acknowledgements

This study is funded by COLIPA.

Correspondence to

Professor Mike Boulton
Department of Ophthalmology & Visual Sciences
The University of Texas Medical Branch
301 University Blvd
Galveston
Texas 77555-1106
USA
e-mail: boultonm@utmb.edu

ICCVAM-NICEATM-ECVAM Symposium on Mechanisms of Chemically-Induced Ocular Injury and Recovery: Current Understanding and Knowledge Gaps

Karen Hamernik[1], Chantra Eskes[2], Jill Merrill[3], Neepa Choksi[4], David Allen[4], James Truax[4], Raymond Tice[5], Marilyn Wind[6], Leonard Schechtman[7] and William Stokes[5]
[1]US EPA*, [2]ECVAM, [3]US FDA, [4]ILS, [5]NICEATM/NIEHS, [6]US CPSC, [7]US FDA,

Summary
A scientific symposium on the Mechanisms of Chemically Induced Ocular Injury and Recovery was held on May 11-12, 2005 in the USA. It was organised and co-sponsored by ICCVAM, NICEATM/NIEHS, and ECVAM, with support from COLIPA. A major goal was to identify research needed to advance the development of test systems that meet regulatory testing requirements and provide for human health protection while reducing, refining (less pain and distress), and/or replacing the use of animals. This paper (one of three), provides a symposium overview and summarises the Session 3 discussions on mechanisms and biomarkers of chemically induced ocular injury and recovery.

Keywords: ocular injury and recovery, chemically-induced, mechanisms of ocular injury and recovery, chemically-related, ICCVAM-NICEATM-ECVAM

Presentation of symposium format and results

A brief overview of the entire symposium is presented in this paper. Also summarised is the part of the meeting (Session 3) dealing with discussions of issues and research needs related to the current understanding of mechanisms of ocular injury and recovery associated with chemical exposure. Other papers by Eskes et al. and Stokes et al., describe, respectively, the other two broad areas covered by the symposium, which were: *In Vitro* Models for Ocular Injury: Current and Potential Biomarkers, and *in Vivo* Models of Ocular Injury and Recovery: Current and Potential Biomarkers to Support Development and Validation of Predictive *in Vitro* Models.

Regulatory need to understand and assess ocular toxicity potential

Accidental eye injury is the leading cause of visual impairment in the United States, and workplace and household chemicals are a significant cause of these injuries, according to sources such as the American Academy of Ophthalmology and the US National Institute for Occupational Safety and Health. US Federal agencies charged with public health protection and the regulation of chemicals or pharmaceuticals and other types of products are concerned about being able to identify potential ocular hazards. Ocular safety and hazard testing in the United States had its origins in part stemming from a well documented case from the 1930s of an eyebrow and eyelash dye that caused severe effects including blindness in women (see illustration on the following US Food and Drug Administration website: http://www.fda.gov/oc/history/historyoffda/section2.html).

Symposium overview

The two-day May 2005 Symposium on the Mechanisms of Chemically Induced Ocular Injury and Recovery was organised and co-sponsored by the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM), the National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM), which is part of the US National Institute of Environmental Health Sciences (NIEHS), and the European Centre for the Validation of Alternative Methods (ECVAM). Additional support was provided by the European Cosmetic, Toiletry and Perfumery Association (COLIPA). The symposium was open to the public and was attended by 76 participants and speakers.

Representatives from ICCVAM, which is composed of 15 US Federal regulatory and/or research agencies, and members of the Ocular Toxicity Working Group (OTWG), one of the numerous ICCVAM subject matter working groups, assisted with symposium organisation, preparation and implementation. Liaisons from ECVAM to the OTWG also provided welcome assistance. A cadre of national and international invited experts from academia, industry, medicine, government, and the animal welfare advocacy area served as speakers and meeting co-chairs.

* The content of this paper does not represent the official position or policy of the US Environmental Protection Agency (US EPA) or any other US federal government agency

One overarching symposium aim was to review the state-of-the-science and current understanding of the pathophysiology, mechanisms and modes of action of chemically induced ocular injury, persistence and recovery. Another aim was to identify areas where research could be conducted to aid the development of test systems that would meet regulatory needs to provide for the protection of human health as well as be compatible with the promotion of animal welfare goals.

Specific symposium objectives were to:

1) Review current and potential molecular, cellular, tissue (e.g. histopathology), and clinical (e.g. corneal opacity, swelling, depth of injury, biomarkers of chemical injury and recovery and their usefulness for *in vivo* and *in vitro* testing models of ocular irritancy and corrosivity);

2) Identify knowledge gaps in the understanding of chemically induced ocular injury and recovery;

3) Identify and prioritise future research initiatives that would address current knowledge gaps and that are considered necessary to advance the development and validation of *in vitro* models of chemically induced ocular injury and recovery; and

4) Discuss and identify quantitative, objective endpoints that should be considered for inclusion in the current *in vivo* rabbit eye test and/or human clinical testing (e.g., more sensitive markers of injury and recovery) that would support development and validation of predictive *in vitro* methods and improve hazard characterisation and reliability.

The symposium agenda over the two-day period consisted of a total of five sessions. During Session 1, an overview of recent initiatives and US regulatory requirements for ocular toxicity testing was provided. Previous workshops on ocular injury and recovery were discussed, including: (1) the International Life Sciences Institute Health and Environmental Sciences Institute (ILSI/HESI) Workshop on Replacing the Draize Eye Irritation Test: Scientific Background and Research Needs, 1995, (Reference: *J. Toxicology – Cutaneous and Ocular Toxicology, 15(3)*, 211-234, 1996); (2) the ILSI/HESI Technical Committee on Alternatives to Animal Testing Expert Meeting on Eye Irritation Testing, Sept 29-30, 1996, (Reference: Ophthalmologic Perspectives on Eye Irritation Testing, Nussenblatt, et al., *J. Toxicology – Cutaneous and Ocular Toxicology 17*,103-109, 1998); (3) the COLIPA Workshop on Mechanisms of Eye Irritation, October 5-8, 1997 (Reference: Report on the COLIPA Workshop on Mechanisms of Eye Irritation, Bruner et al., *ATLA 26*, 811-820, 1998). Also mentioned was the ECVAM Workshop on Eye Irritation Testing: The Way Forward (Reference: Balls et al., *ATLA 27*, 53-77, 1998). Also noteworthy, are the results of two international workshops organised and sponsored by the [US] Interagency Regulatory Alternatives Group (IRAG) in 1991 and 1993, respectively: (1) Workshop on Updating Eye Irritation Test Methods: Proposals for Regulatory Consensus (Reference: *Food Chemical Toxicol. 31(2)*, 1993) and (2) Workshop on Eye Irritation Testing: Practical Applications of Non-Whole Animal Alternatives (Reference: *Food Chemical Toxicol. 35 (1)*, 1997).

Session 2 of the symposium reviewed current ocular injury and toxicity assessments and included a discussion of human chemically induced ocular injury. Session 3 (summarised in this paper) dealt with mechanisms and biomarkers of ocular injury and recovery. In Session 4, current *in vitro* models of ocular injury and recovery were discussed. The subject of Session 5 was *in vivo* quantitative objective endpoints to support development and validation of predictive *in vitro* models. Panel discussions were held after Sessions 3, 4 and 5. Each of the panels was given a set of pertinent questions to address based on the session topic, and the discussions were summarised at the end of the meeting.

Mechanisms and biomarkers of ocular injury and recovery

Symposium Session 3 focused on issues related to the current understanding of known mechanisms and modes of action of chemically related ocular injury, persistence and recovery. Speakers in this session discussed topics such as eye injury type and reversibility, cellular (e.g. corneal epithelium and endothelium) and tissue (e.g. cornea, conjunctiva, iris) responses to chemical injury in humans and animals, chemical-specific modes of action (e.g. for acids and bases) and the role of histopathology and depth of injury in evaluating ocular injury onset, extent, severity and recovery potential. Other speakers presented information on (1) the role of chemical toxicokinetics in ocular injury (including delayed onset of effect) and detoxification, (2) possible future roles for toxicogenomics in elucidating processes involved in ocular injury and its sequelae, and (3) the effects of chemical exposure on tear film.

Session 3 panel piscussion questions and responses

The panel for Session 3 was asked to respond to a series of four questions. Highlights of the panel's answers and opinions are presented below.

In the first question, the panel was asked to identify the currently known mechanisms and modes of action of chemically induced ocular injury and recovery. The panel responded that mechanisms of injury (e.g. cytotoxicity, protein coagulation, membrane saponification, disruption of the extracellular matrix, inflammatory cell infiltration with release of mediators, up-regulation of proteases and collagenases) are known for some chemicals and product types.

Further, the extent of ocular surface involvement and depth of corneal penetration may correlate with severity of lesions and recovery and thus could serve as a biomarker of reversibility potential. The panel proposed that more research could provide information on the general applicability of this concept. In addition, it was noted that some existing studies have shown species differences in response to the same chemical and dose-dependent differences in response to the same chemical (i.e. mechanisms of repair or injury at lower doses may not be relevant at higher doses). Currently, more data and information are available on severe ocular injury than for milder forms.

For the second question, the panel was asked to list current knowledge gaps in understanding of mechanisms and modes of

action of chemically induced ocular injuries and recovery. The panel responded that further assessment was needed of the relationship between type and severity of initial *in vivo* damage and persistence of effects. Furthermore, additional work was required to identify and develop for utility, biomarkers of injury and recovery (i.e. gene expression profiling, clustering and pathway analysis for ocular damage and repair) and to elucidate the role of tear film in ocular damage. It also would be important to improve the translation of *in vivo* physiology to *in vitro* models. This would involve enhancing current knowledge of such things as the role of metabolism and the linkage of expression of specific corneal proteins with injury and recovery processes.

In question three, the panel was asked to identify research initiatives needed to address current knowledge gaps and further characterise mechanisms and modes of action in order to advance the development and validation of predictive *in vitro* models of chemically induced ocular injury and recovery. The panel's opinion was that quantitative endpoints could be incorporated into the current *in vivo* test and evaluated to make the test more informative. These would include histopathology to correlate cellular changes with observational endpoints, HPLC and mass spectroscopy to evaluate penetration of substances into the eye, and depth of injury analysis to gain further insight on the utility of this measurement as a biomarker for reversibility/irreversibility of effects. The panel thought that further evaluation of species differences and dose-dependent differences in response was merited. Also mentioned was the need for additional work on tear film, especially with regard to composition determination, further elucidation of its role in protection of ocular constituents and the potential consequences of its disruption on the severity of effects of mild and moderate irritants. Other areas for additional work included the need to better characterise early onset versus delayed ocular responses to chemical agents, to investigate further the toxicokinetics of chemical exposure to the eye, to better elucidate the role of inflammatory responses in observed ocular damage and to better explore recovery mechanisms of the eye (such as effects on stem cells). Also suggested was additional evaluation of other *in vitro* models that might be more predictive or useful in hazard identification, such as human corneal models (isolated and reconstituted) and pig corneal models.

For the final question, the panel was asked to identify what *in vivo* biomarkers (e.g. molecular, cellular, morphological, clinical) should be further investigated as predictive indicators of severity of lesions, reversibility versus non-reversibility, or delayed responses. In answer, the panel mentioned as noteworthy: histopathology, quantitative endpoints obtained using standard biomicroscopy, confocal microscopy, selective staining, cytology, immunologic markers and gene expression profiling, and clustering and pathway analysis for ocular damage and repair.

A report of the Symposium proceedings will be available in the near future on the ICCVAM website (http://iccvam.niehs.nih.gov).

Correspondence to
Dr. Karen Hamernik (WHO Temporary Adviser)
Office of Science Coordination and Policy
US Environmental Protection Agency (EPA)
EPA East Building (Mail Code 7501M)
1200 Pennsylvania Avenue, NW
Washington, DC 20460
USA
e-mail: hamernik.karen@epamail.epa.gov

An Overview of the COLIPA Eye Irritation Research Programme

Penny Jones[1], Dan Bagley[2], Claudine Faller[3], Beatrice Le Varlet[4], Pauline McNamee[5], Irene Manou[6], Wolfgang Pape[7], Christine Van den Berghe[8] and Freddy Van Goethem[9]

[1]Unilever-Safety and Environmental Assurance Centre, Colworth Park, Sharnbrook, Bedford, UK; [2]Colgate-Palmolive, Piscataway, NJ, USA; [3]Cosmital Wella, Marly, Switzerland; [4]LVMH Recherche, Saint Jean de Braye, France; [5]The Procter & Gamble Company, Egham, Surrey, UK; [6]COLIPA, Brussels, Belgium; [7]Beiersdorf, Hamburg, Germany; [8]L'Oreal, Aulnay Sous Bois, France; [9]Johnson & Johnson, Pharmaceutical Research & Development, Beerse, Belgium

Summary

The COLIPA eye irritation programme incorporates research projects and collaborative activities with external partners. Three projects focus on understanding mechanisms of eye injury and identification of new in vitro endpoints that are more predictive of the in vivo human response to chemical injury: 1) investigation of whether kinetics/patterns of change in physiological function and signals of injury released from the cornea in vitro can predict a chemical's potential to damage the eye, with a focus on recovery; 2) identification of endpoints related to the magnitude of injury and quality of repair in human immortalised cells and 3D human conjunctival and corneal constructs; 3) a genomics project using a pattern recognition approach to identify new endpoints for injury/repair that builds on corneal models from projects 1 and 2. Collaboration of industry, academia, external scientific organisations and regulators is equally important. COLIPA is working with ECVAM by actively participating in its Eye Irritation Task Force and providing support for statistical analysis of current in vitro methods.

Keywords: in vitro, eye irritation, Draize eye test, alternative, replacement, COLIPA

Introduction

The efforts of the COLIPA (European Cosmetic, Toiletry and Perfumery Association) -SCAAT (Steering Committee on Alternatives to Animal Testing) Eye Irritation Task Force are dedicated to the development of *in vitro* alternative methods/strategies to assess eye irritation for the replacement of the Draize test. The research programme follows on from previous validation studies (e.g., Bagley et al., 1992; Balls et al., 1995; Brantom et al., 1997) and workshops (Bruner et al., 1998; Balls et al., 1999). The approach used incorporates integrated research projects and collaborative activities with external partners.

Research programme strategy and objectives

The approach of the COLIPA eye irritation research programme is to build on the experience of the earlier validation studies and scientific workshops through collaborations with academic institutions conducting COLIPA-supported projects and other collaborative activities with external partners. The objective is to gain an understanding of cellular molecular mechanisms of chemically induced eye irritation, with focus on corneal injury and recovery. Through this understanding the expected outcome is the identification of *in vitro* endpoints related to the dynamics of injury and recovery that are more predictive of the *in vivo* human response to chemical injury. This will enable the development of prediction models for pre-validation of new or improved *in vitro* methods that would proceed to formal validation.

There are three integrated research projects: 1) an investigation of whether kinetics/patterns of change in physiological function and signals of injury released from the cornea *in vitro* can predict a chemical's potential to damage the eye, with a focus on recovery; 2) identification of endpoints related to magnitude of injury and quality of repair in human immortalised cells and 3-dimensional human conjunctival and corneal constructs and 3) a genomics project using a pattern recognition approach to identify new endpoints for injury and repair that builds on the corneal models being evaluated in projects 1 and 2 for potential use in current/future *in vitro* assays. The approaches taken for each of these projects are given below.

Research Projects

Project 1: *In vitro* corneal culture eye irritation assay

This project is being conducted by Norbert Schrage and Markus Frenz at the University of Aachen, Germany, and was initiated in January 2002. The aims are 1) to develop an *in vitro* model of excised corneas maintained in culture to allow observation of injury and recovery following chemical exposure and 2) to investigate whether kinetics/patterns of change in physiological function and signals of injury released from the perfused cornea *in vitro* can predict a chemical's potential to damage the eye, with a focus on recovery.

The following stepwise approach has been adopted to develop a new isolated perfused corneal culture model that can be maintained for a period of time under steady culture conditions:

- Determination of viability and stability of the isolated perfused corneal culture system both morphologically and metabolically, and definition of the parameters to be used routinely to confirm system viability and stability
- Determination of suitability of the model to investigate wound healing by mechanical abrasion
- Exposure of the defined isolated perfused corneal system to model toxicants
- Identification of the morphological/biochemical markers of injury/recovery to be used routinely to evaluate toxicant effects
- Investigation of evaluation methods including biomicroscopy, pachymetry and glucose/lactate turnover for system viability/stability, LDH, cytokines (IL-1α, IL-2, IL-6 IL-8, MIP1) and growth factors (FGF, VEGF) for evaluation of dynamics of injury and recovery after mechanical trauma or toxicant exposure.

Project 2: Cell culture models for ocular toxicity studies

This project is being undertaken by Monica Berry and Marcus Radburn-Smith at the University of Bristol, UK, and was initiated in January 2002. The aims of this project are 1) to sequentially build 3D human corneal constructs consisting of epithelium, stroma and endothelium in order to better understand underlying mechanisms of action of eye irritation and 2) to identify new endpoints related to magnitude of injury and quality of repair in human corneal models that will enable prediction of the nature and severity of toxicant effects.

This following approach is therefore being used in this project to investigate the physiological responses to ocular injury (e.g. cell activation/signalling to immune system effector cells) by evaluating responses to model toxicants in increasingly complex corneal constructs:

- Investigation of human corneal and conjunctival cell lines for culture conditions, growth characteristics and suitability for use in 3D constructs
- Development of stratified epithelia, stromal equivalents and construction of two layer models (epithelium and stroma)
- Construction of three layer models by the addition of an endothelium cell layer
- Exposure of monolayers and stratified models to model toxicants using the following evaluation methods
- Light and confocal microscopy, characterisation of surface markers and differentiation, barrier formation assessment, membrane damage, metabolic activity, profiling of cytokine secretion.

Project 3: Development of gene expression fingerprints to identify toxic damage to the cornea

This project will be initiated in the summer of 2005 and will be led by Mike Boulton, University of Cardiff, U.K. The aims of this project are 1) to generate proof of concept that generic chemicals will cause differential gene expression in human bioengineered corneas 2) to identify gene expression profiles in bioengineered human corneas exposed to generic classes of chemicals and 3) to develop a gene fingerprint directory to identify chemicals toxic to a bioengineered human cornea. The principal outcome should be the application of the knowledge to better understand new endpoints for eye irritation and enable further development of current and future *in vitro* methods.

This project will use the following approach:

- Microarray analysis with human Affymetrix Gene Array chips using RNA from untreated versus treated (model toxicants) bioengineered human corneal constructs developed in the Bristol project
- Analysis expansion to define a gene expression fingerprint for the effects of each chemical class. Fingerprints will be generated for two time points to assess both injury and repair.

External collaboration

Equally important to achieve validated *in vitro* methods is collaboration between industry, academia, external scientific organisations and regulators. COLIPA is working with ECVAM by active mutual participation in both COLIPA and ECVAM Eye Irritation Task Forces to ensure that the research efforts are synergistic rather than duplicative between the two organisations. The Task Force also uses external contacts in order to benchmark and refine the eye irritation research projects.

References

Bagley, D. M., Bruner, L. H., de Silva, O. et al. (1992). An evaluation of five potential alternatives in vitro to the rabbit eye irritation test in vivo. *Toxicology in Vitro 6*, 275-284.

Balls, M,, Botham, P. A., Bruner, L. H. and Spielmann, H. (1995). The EC/HO international validation study on alternatives to the Draize eye irritation test. *Toxicology In Vitro 9*, 871-929.

Balls, M., Berg, N., Bruner, L. H. et al. (1999). The report and recommendations of ECVAM workshop 34. *ATLA 27*, 53-77.

Brantom, P. G., Bruner, L. H., Chamberlain, M. et al. (1997). A summary report of the COLIPA international validation study on alternatives to the Draize rabbit eye irritation test. *Toxicology in Vitro 11*, 141-179.

Bruner, L. H., de Silva, O., Earl, L. K. et al. (1998). Report on the COLIPA workshop on mechanisms of eye irritation. *ATLA 26*, 811-820.

Acknowledgments

The Eye Irritation Taskforce members would like to acknowledge the support of COLIPA-SCAAT for this research programme and the collaborating researchers at the Universities of Aachen, Bristol and Cardiff.

Correspondence to

Penny Jones
Safety and Environmental Assurance Centre
Colworth Park
Sharnbrook, Bedford UK
e-mail: Penny.Jones@unilever.com

Workshop 5.12
Toxicogenomics – potential, validation and case studies

Development of an *In Vitro* Gene Expression Assay for Predicting Hepatotoxicity

Stephen Abel, Yi Yang and Jeffrey F. Waring
Abbott Laboratories, Dept. of Cellular and Molecular Toxicology, Abbott Park, IL, USA

Summary
A significant number of compounds do not proceed in drug discovery due to toxicity issues. The purpose of this study was to evaluate whether gene expression profiles could be identified and used to classify drugs based on the mechanism of toxicity in an in vitro system. Rat hepatocytes were treated with two classes of drugs, aromatic hydrocarbon receptor (AhR) and peroxisome proliferator activated receptor (PPAR) ligands.
The results showed that a small set of genes could be identified that could be used to classify new compounds with distinct mechanisms of toxicity.

Keywords: toxicogenomics, PPAR, AhR, primary rat hepatocytes

Introduction

Toxicology studies are a major bottleneck in the development of new drug candidates (Service, 2004). One of the principle reasons for this is that the majority of toxicology studies are conducted in animals, requiring large amounts of compounds and significant resources (Kola and Landis, 2004). The development of an assay that could screen new drug candidates *in vitro* for toxicity before animal studies are conducted would result in significant savings in time and resources, and would result in fewer compounds with toxic liabilities coming through the discovery process.

Toxicogenomics is the application of gene expression analysis towards drug safety evaluation (Guerreiro et al., 2003). Currently, *in vivo* toxicogenomic studies have been used to show that drugs with similar mechanisms of toxicity produce characteristic gene profiles and to study mechanisms of toxicity (Jun et al., 2004; Waring et al., 2002). Toxicogenomics has also been applied in a predictive mode where gene expression changes can be noted before animals show changes in clinical pathology or histopathologic lesions (Ellinger-Ziegelbaur et al., 2004; Kier et al., 2004). Recent publications have applied toxicogenomics using *in vitro* systems to study toxic mechanisms and identify gene expression changes associated with certain mechanisms of toxicity (Harris et al., 2004; Sawada et al., 2005). In this study, toxicogenomics was used to determine a characteristic gene expression profile for a class of compounds, referred to as a "gene signature".

The objective of this study was to develop *in vitro* gene signatures for two relevant, drug-induced toxicities: Aryl hydrocarbon receptor (AhR) and peroxisome proliferator activated receptor (PPAR). PPARα and AhR proteins are nuclear receptors that act as transcription factors when bound to appropriate ligands (Jacobs et al., 2003). Pathology associated with PPARα binding includes an increase in peroxisomes as well as tumour development (Cattley and Roberts, 2000). AhR activation induces hepatotoxicity, immune suppression and carcinogenicity (Nebert et al., 2004; Poland and Knutson, 1982). Results from our studies show that *in vitro* gene signatures can be generated from rat hepatocytes treated with reference prototypical PPAR and AhR ligands and that these signatures can be used for *in vitro* characterisation of experimental compounds.

Animals, materials, methods

Culture conditions

Primary rat hepatocytes were obtained as previously described (Waring et al., 2002). Rat hepatocytes were plated at a density of 800,000 cells/ml with 10 ml InVitroGro™ Hepatocyte Medium (In Vitro Technologies, Baltimore, MD, USA) on 100 mm BD BioCoat™ Collagen I Cellware (Becton Dickinson, Bedford, MA, USA) at 37°C and 5% CO_2 . Hepatocytes were treated 48 hours with a TC10 dose for the three drugs of each class: PPAR (WY-14643, bezafibrate, clofibrate, troglitazone, fenoprofen, indomethacin), AhR (3-methylcholanthrene, beta-napthoflavone, aroclor, omeprazole, benzopyrene, A277249, 3-3' diindolylmethane) and negative controls (chlorpheniramine, penicillin, spectinomycin).

Drug sources

The drugs were sourced from Sigma Laboratories with the following exceptions: A-277249 was obtained from Abbott Laboratories, 3,3'diindolylmethane (DIM) was obtained from A Better Choice for Research Chemicals (ABCR, Karlsruhe, Germany), and aroclor was obtained from Monsanto (St. Louis, MO, USA).

RNA isolation

RNA was isolated from hepatocytes using TRIzol® (Invitrogen Life Technologies, Carlsbad, CA, USA) reagent according to the manufacturer's protocol. The RNA concentration was determined by absorbance at 260 nm with the Smart-Spec spectrophotometer (Bio-Rad, Hercules, CA, USA). RNA integrity was evaluated using the RNA 6000 Nano Assay with the 2100 Agilent bioanalyzer (Agilent Technologies, GmbH, Germany) according to the manufacturer's protocol.

MTT cell cytotoxicity assay

The colorimetric cell viability assay using the tetrazolium salt (3-(4,5-dimethylthiazol-2-yl)-2,5,-diphenyl tetrazolium bromide (MTT) was used as previously described (Mossmann, 1983).

Microarrays

Microarray analysis was performed on Affymetrix RAE-230A chips according to a protocol provided by Affymetrix (Affymetrix, Santa Clara, CA, USA). The manufacturer's protocol was followed with the exception that the primer used for the reverse transcription reaction was the GeneChip® T7-Oligo(dT) Promoter Primer Kit (Affymetrix, Santa Clara, CA, USA).

Bioinformatics

The intensity data from the scanned image of the microarray were imported into Rosetta Resolver® gene expression analysis software version 4.0 (Rosetta Biosoftware, Seattle, WA, USA). Intensity based ratios were calculated in Resolver using intensity data from each treatment microarray versus its respective vehicle control. Two vehicle treated samples were pooled *in silico* for each ratio. Hierarchical clustering was done in Resolver using the agglomerative algorithm and Euclidean distance for finding significant genes with expression ratios ≥2.0 and p-values ≤0.01.

Gene signature algorithm

Each drug was tested in duplicate across three separate rat hepatocyte isolations. An experiment was defined as the data from a single microarray. A total of 18 experiments were performed for each class (3 compounds per class x 2 technical replicates x 3 rat hepatocyte isolations). Experiments were divided into two groups: training and verification. The training experiments were used to determine the gene signatures, and the verification experiments were used to verify if the signatures could be used

Tab. 1: MTT cytotoxicity data for compounds tested. Treatment dose set at 10% cytotoxicity.

Drug Classification	Drug	Treatment Dose [µM]	Gene Signature Training Set	Gene Signature Verification
Peroxisome Proliferation	Wy-14643	35	X	
	Bezafibrate	200	X	
	Clofibrate	550	X	
	Troglitazone	75		X
	Fenoprofen	300		X
	Indomethacin	235		X
AhR Ligand	3-methylcholanthrene	100	X	
	Beta-napthoflavone	75	X	
	Aroclor	100	X	
	Omeprazole	300		X
	Benzopyrene	10		X
	A-277249	108		X
	3,3' diindolylmethane (DIM)	75		X
Negative Control	Chlorpheniramine	275	X	
	Penicillin	150	X	
	Spectinomycin	90	X	

to correctly identify drugs by class. Drugs for verification were tested in duplicate with a single rat hepatocyte isolation.

The first step of the algorithm was to find significantly regulated genes for each drug of a class from among the approximately 15,000 genes represented on the microarray. Parameters were set to select genes that displayed at least a two-fold change (either up or down) with a p-value ≤0.01. Genes that were commonly regulated among the three drugs were then selected for each class. The gene list was then narrowed to include genes significantly regulated in at least 16 of the 18 total experiments within a class.

Significantly regulated genes from the negative controls were defined as those with fold-changes ≥2.0 with a p-value ≤0.01 in at least six of the negative control experiments. These genes were then removed from the PPAR or AhR gene signature lists.

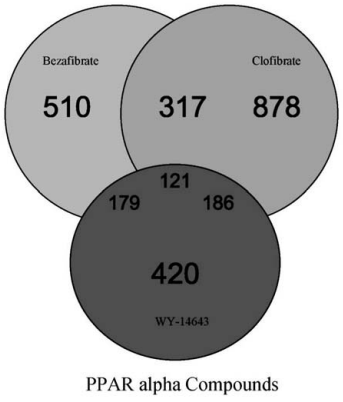


Fig. 1: Venn diagram showing that there are 121 commonly regulated genes for the PPARα training set compounds. The coloured figures can be downloaded from the ALTEX website (www.altex.ch).

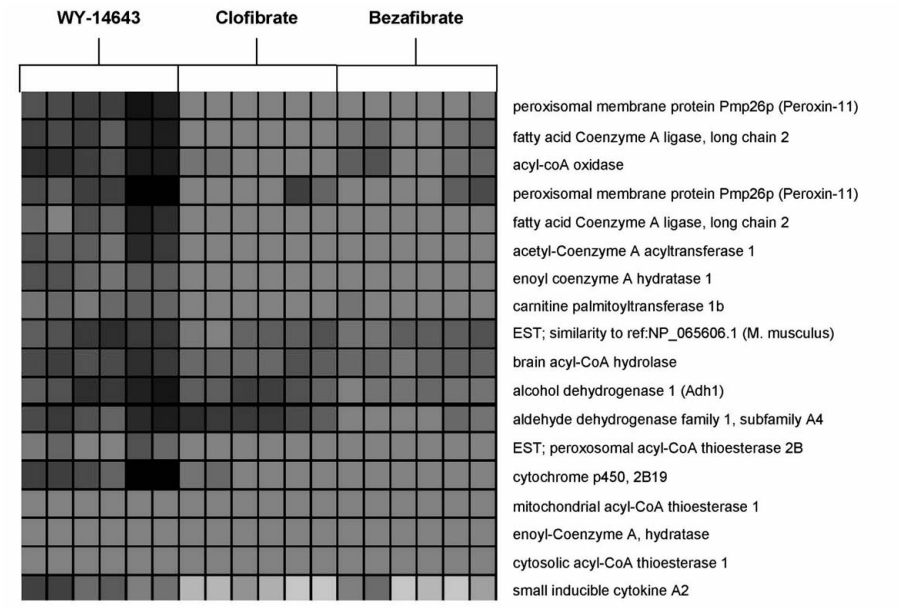


Fig. 2: Ordered cluster analysis showing gene expression for the 18 genes of the PPAR alpha gene signature. Data shown as a fold change relative to vehicle (1% DMSO), p value ≤0.01, with a range of five-fold reduced expression (light grey) to five-fold increased expression (dark grey).

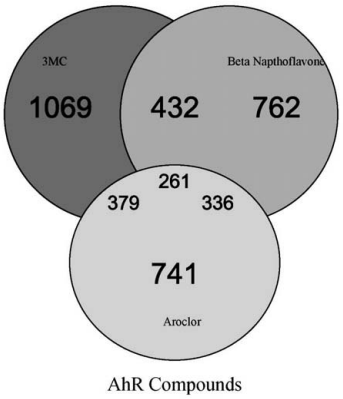


Fig. 3: Venn diagram showing that there are 261 commonly regulated genes for AhR training set compounds.

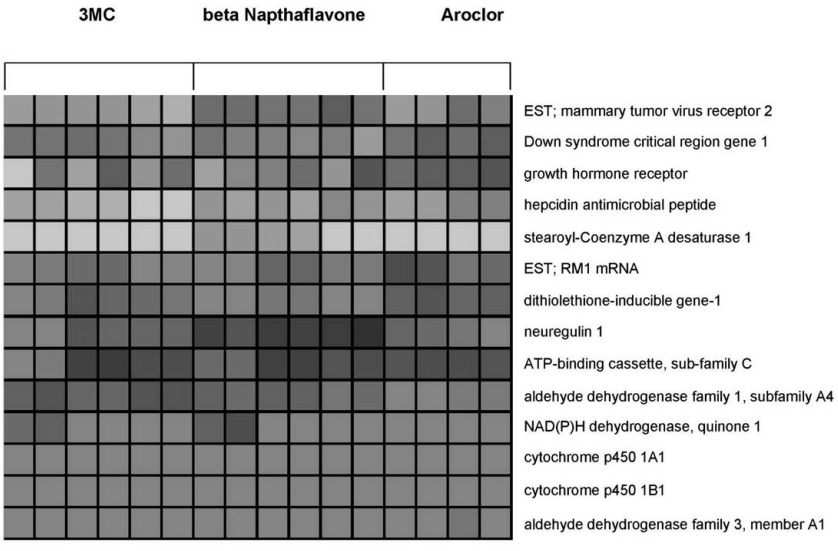


Fig. 4: Ordered cluster analysis showing gene expression for 14 genes of the AhR gene signature. Data shown as fold change relative to vehicle (1% DMSO), p ≤0.01, with a range of five-fold reduced expression (light grey) to five-fold increased expression (dark grey).

Results

Cell cytotoxicity assay

The MTT cytotoxicity results shown in table 1 are the doses calculated to result in 90% survival (10% cytotoxicity). The exceptions are benzopyrene, which was tested at 10 µM due to solubility limitations, and DIM, which was tested at 75 µM, as previously reported (Hestermann and Brown, 2003). All drugs were tested in solutions containing less than 1% DMSO vehicle.

Gene signature algorithm

Data for the PPAR training compounds are shown in figure 1 as a Venn diagram. The PPAR class contained the following significantly regulated genes across three separate isolations: 578 for bezafibrate, 878 for clofibrate, and 420 for Wy-14643. A total of 121 genes were commonly regulated by all three drugs. From the 121 common genes, 27 genes were found to be consistently regulated across the three isolations. Genes in common with negative controls were removed and the 18 remaining genes are listed in figure 2, with the experiments shown in an ordered cluster analysis. The genes for peroxisomal membrane protein and fatty acid Coenzyme A ligase are listed twice because these genes are represented on the microarray by two different probe sequences.

Data for the AhR training compounds are shown in figure 3 as a Venn diagram. The AhR class contained the following significantly regulated genes: 1069 for 3MC, 762 for beta-napthoflavone, and 741 for aroclor. A total of 261 genes were commonly regulated by all three drugs. Of the 261 common genes, 37 genes were found to be consistently regulated across the three isolations. Genes in common with negative controls were removed and the 14 remaining genes are listed in figure 4, with the experiments shown in an ordered cluster analysis. The data from two of the six aroclor experiments are not shown due to RNA degradation.

Gene signature verification

The gene signatures were verified with additional reference compounds from the same class. The hierarchical cluster analysis shown in figure 5 shows that benzopyrene, A277249, omeprazole, and DIM all cluster with the AhR training compounds and not with the negative control or the PPAR compounds. Likewise, the PPARα verification compounds indo-methacin and fenoprofen) cluster with PPARα training compounds. In contrast, troglitazone clusters with the negative control compounds and not the PPARα compounds. Finally, the AhR, PPAR, and negative control compounds all cluster separate from each other, with each class of compound producing a

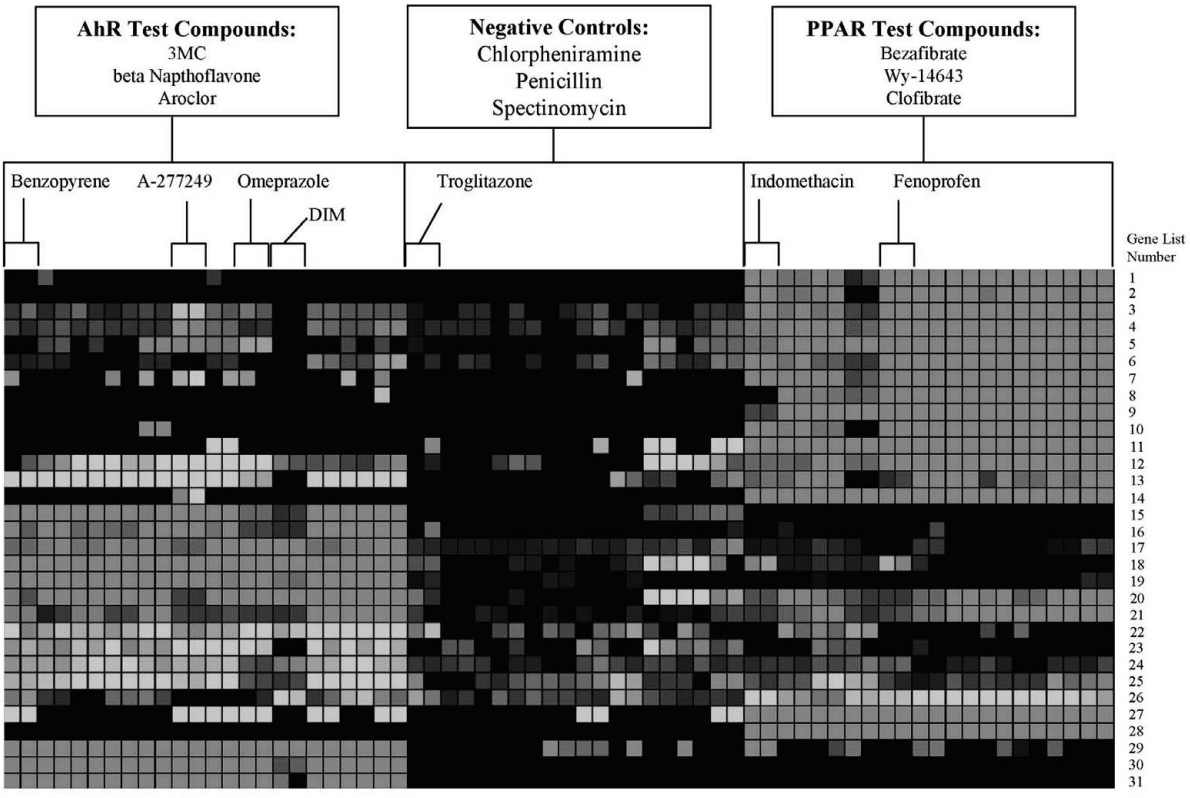


Fig. 5: Hierarchical cluster analysis showing gene expression for all experiments using the PPARα and AhR gene signatures. Data shown as fold change relative to vehicle (1% DMSO), p ≤0.01, with a range of five-fold reduced expression (light grey) to five- fold increased expression (dark grey).

unique gene expression signature. The corresponding genes are presented in order in table 2. Note that 31 genes are listed because both signatures share the gene, aldehyde dehydrogenase family 1, subfamily A4.

Discussion

The objective of this study was to develop gene signatures for two well characterised classes of compounds, PPARα and AhR agonists, and to use these signatures to correctly identify other compounds of the same class using an *in vitro* rat hepatocyte system.

Gene signatures were developed for the PPARα and AhR pathways by treating primary rat hepatocytes with class-specific compounds and measuring gene expression changes with rat DNA microarrays. The compounds were tested across three different rat hepatocyte isolations to capture biological variability and in duplicate within each isolation to capture technical variability. Three reference compounds were selected for each class and used as a training dataset to determine a gene signature. The negative control compounds were used in an attempt to identify gene expression changes that were due in part to the *in vitro* conditions. 18 genes were identified for the PPARα signature and 14 genes for the AhR signature.

Tab. 2: PPAR and AhR gene list corresponding to data in figure 5.

Gene List Number	Gene Name
1	peroxisomal membrane protein Pmp26p (Peroxin-11)
2	peroxisomal membrane protein Pmp26p (Peroxin-11)
3	fatty acid Coenzyme A ligase, long chain 2
4	acetyl-Coenzyme A acyltransferase 1
5	enoyl-Coenzyme A, hydratase
6	acyl-coA oxidase
7	fatty acid Coenzyme A ligase, long chain 2
8	brain acyl-CoA hydrolase
9	carnitine palmitoyltransferase 1b
10	cytochrome p450, 2B19
11	EST; peroxosomal acyl-CoA thioesterase 2B
12	EST; similarity to ref:NP_065606.1 (M. musculus)
13	stearoyl-Coenzyme A desaturase 1
14	mitochondrial acyl-CoA thioesterase 1
15	EST; RM1 mRNA
16	neuregulin 1
17	ATP-binding cassette, sub-family C
18	dithiolethione-inducible gene-1
19	NAD(P)H dehydrogenase, quinone 1
20	aldehyde dehydrogenase family 1, subfamily A4
21	alcohol dehydrogenase 1 (Adh1)
22	EST; mammary tumor virus receptor 2
23	growth hormone receptor
24	Down syndrome critical region gene 1
25	hepcidin antimicrobial peptide
26	small inducible cytokine A2
27	enoyl-Coenzyme A, hydratase
28	cytosolic acyl-CoA thioesterase 1
29	cytochrome p450 1A1
30	cytochrome p450 1B1
31	aldehyde dehydrogenase family 3, member A1

The gene signatures were verified using additional compounds from each class. Both omeprazole and DIM are considered to be weak non-classical AhR agonists without the structural features of AhR ligands (Hestermann and Brown, 2003; Backlund et al., 1997). The results here indicate that the gene signature is sensitive enough to identify a range of structurally diverse drugs with similar biological properties (Hestermann and Brown, 2003; Backlund et al., 1997). Troglitazone, a PPARγ agonist, clustered with the negative controls. PPARγ receptors are expressed primarily in adipose tissue and affect insulin and lipid metabolic pathways (Lebovitz and Banerji, 2001). These data suggest that the PPAR gene signature is specific for the α class of receptor and can differentiate between the PPARα and PPARγ pathways.

The genes listed in the PPAR and AhR signatures correlate with the biology of the receptors. PPAR genes associated with β-oxidation regulated in *in vivo* studies include acyl-CoA oxidase, acyl-CoA thioesterase, enoyl-CoA hydratase, fatty acid coenzyme A ligase, and aldehyde dehydrogenase (Baker et al., 2004). The genes regulated for AhR ligands include CYP1A, CYP1B, NAD(P)H reductase, and aldehyde dehydrogenase and have been previously reported to be regulated in *in vivo* studies (Waring et. al., 2001). Genes missing from the signatures include CYP4A for PPAR and glutathione-s-transferase (GST) for AhR. These genes were removed from their corresponding signature due to regulation by negative controls and inconsistent regulation across the three isolations (data not shown).

The PPAR signature was also tested against the *in vitro* conditions of drug dose and treatment time. Hepatocytes were treated with clofibrate, WY-14643, and bezafibrate for 4, 24, 48 and 72 hours. The dose for clofibrate was ten times lower than previously tested (55 μM). The low dose clofibrate treatment clustered with the negative control compounds at all time points (data not shown). The WY-14643 and bezafibrate four-hour treatment clustered with the negative controls, whereas the 24, 48, and 72-hour treatments clustered with the other PPAR compounds (data not shown). These data indicate that the signature is sensitive to dosing as well as to treatment time, with 24 hours giving similar results to 48 and 72-hour treatment times, suggesting that 24 hours is adequate for treatment.

In summary, this study demonstrates that an *in vitro* assay using primary rat hepatocytes and gene expression analysis can successfully be used to predict hepatotoxicity. An assay of this type will have significant impact on the drug discovery process by allowing drugs to be screened prior to animal studies and obtaining toxicology information earlier in the development process. By filling these needs, this *in vitro* assay will help to determine which drugs offer the best chance for success.

References

Backlund, M., Hohansson, I., Mkrtchian, S. and Ingelman-Sundberg, M. (1997). Signal transduction-mediated activation of the aryl hydrocarbon receptor in rat hepatoma H4IIE cells. *Journal of Biological Chemistry 272 (50)*, 31755-31763.

Baker, V. A., Harries, H. M., Waring, J. W. et al. (2004). Clofibrate-induced gene expression changes in rat liver: a cross-laboratory analysis using membrane cDNA arrays. *Environmental Health Perspectives 112 (4)*, 428-438.

Cattley, R. C. and Roberts, R. (2000). Peroxisome proliferators and carcinogenesis: editorial perspectives. *Mutation Research 448*, 117-119.

Ellinger-Ziegelbauer, H., Stuart, B., Wahle, B. et al. (2004). Characteristic expression profiles induced by genotoxic carcinogens in rat liver. *Toxicological Science 77 (1)*, 19-34.

Guerreiro, N., Staedtler, F., Grenet, O. et al. (2003). Toxicogenomics in Drug Development. *Toxicologic Pathology 31*, 471-479.

Harris, A. J., Dial, S. L. and Casciano, D. A. (2004). Comparison of basal gene expression profiles and effects of hepatocarcinogens on gene expression in cultured primary human hepatocytes and HepG2 cells. *Mutation Research May 18, 549 (1-2)*, 79-99.

Hestermann, E. V. and Brown, M. (2003). Agonist and chemopreventative ligands induce differential transcriptional cofactor recruitment by aryl hydrocarbon receptor. *Molecular Cell Biology 23*, 7920-7925.

Jacobs, M. N., Dickins, M. and Lewis, D. F. (2003). Homology modelling of nuclear receptors: human oestrogen receptor beta (hERbeta), the human pregnane-X receptor (PXR), the Ah receptor (AhR) and the constitutive andorstane receptor (CAR) ligand binding domains from the human oestrogen receptor alpha (hERalpha) crystal structure, and the human peroxisome proliferator activated receptor alpha (PPARalpha) ligand binding domain from the human PPARgamma crystal structure. *Journal Steroid Biochemistry Molecular Biology Feb, 84 (2-3)*, 117-132.

Jun, J. W., Park, J. S., Hwang, J. W. et al. (2004). Gene expression analysis of peroxisome proliferators-and phenytoin-induced hepatotoxicity using cDNA microarrays. *Journal of Veterinary Medical Science Nov 66(11)*, 1329-1333.

Kier, L. D., Neft, R., Tang, L. et al. (2004). Applications of microarrays with toxicologically –relevant genes (tox genes) for the evaluation of chemical toxicants in Sprague Dawley rats in vivo and human hepatocytes. *Mutation Research May 18, 549 (1-2)*, 101-113.

Kola I. and Landis J. (2004). Can the pharmaceutical industry reduce attrition rates? *Nature Reviews Drug Discovery 3(8)*, 711-715.

Lebovitz, H. E. and Banerji, M. A. (2001). Insulin resistance and its treatment by thiazolidinediones. *Recent Progress in Hormone Research 56*, 265-294.

Mossmann, T. (1983). Rapid colorimetric assay for cellular growth and survival: applications to proliferation and cytotoxicity assays. *Journal of Immunological Methods 65*, 55-63.

Nebert, D. W., Dalton, T. P., Okey, A. B. and Gonzalez, F. J. (2004). Role of aryl hydrocarbon receptor-mediated induction of the CYP1 enzymes in environmental toxicity and cancer. *Journal of Biological Chemistry 279(23)*, 23847-23850.

Poland, A. and Knutson, J. C. (1982). 2,3,7,8-Tetra-chlorodibenzo-p-dioxin and related halogenated aromatic hydrocarbons: examination of the mechanism of toxicity. *Annual Review Pharmocology Toxicology 22*, 517-554.

Sawada, H., Takami, K. and Asahi, S. (2005). A toxicogenomic approach to drug-induced phospholipidosis: analysis of its induction mechanism and establishment of a novel in vitro screening system. *Toxicological Sciences February, 83 (2)*, 282-292.

Service, R. F. (2004). Surviving the blockbuster syndrome. *Science Mar 19; 303(5665)*, 1796-1799.

Suolinna, E. M. (1982). Isolation and culture of liver cells and their use in biochemical research of xenobiotics. *Medical Biology Oct 60 (5)*, 237-254.

Waring, J. F., Gum, R., Morfitt, D. et al. (2002). Identifying toxic mechanisms using DNA microarrays: evidence that an experimental inhibitor of cell adhesion molecule expression signals through the aryl hydrocarbon nuclear receptor. *Toxicology 181-182*, 537-550.

Waring, J. F., Jolly, R. A., Ciurlionis, R. et al. (2001). Clustering of hepatotoxins based on mechanism of toxicity using gene expression profiles. *Toxicology Applied Pharmacology 175*, 28-42.

Correspondence to

Dr Jeffrey F. Waring
Dept R463, AP9A
100 Abbott Park Road
Abbott Park, IL 60064
USA
e-mail: Jeff.Waring@abbott.com

Workshop 5.13
Strategies for prioritising and streamlining the validation process

Validation *via* Weight-of-Evidence Approaches

Michael Balls and Robert Combes
FRAME, Nottingham, UK

Summary
It is not always necessary, or even possible, to conduct a practical laboratory study to establish the validity of tests or testing strategies. A weight-of-evidence approach aims to use already-available information in a structured, systematic, independent and transparent assessment. Crucial aspects include: study management and design; selection of experts without vested interests; collection of data and data quality control; differential weighing of various types of evidence, alone and in combination; evaluation of test performance in terms of reliability and relevance in relation to purpose; publication of outcome; and peer review.

Keywords: invalidation, test, validation, weight-of-evidence

Introduction

Early discussions on evaluating the reliability and relevance of non-animal tests for particular purposes, i.e. on their validation, focused on the performance of practical multi-laboratory studies (Balls et al., 1990, 1995). However, it subsequently became clear that such practical studies would not always be necessary or even possible, and that, in some circumstances, a weight-of-evidence approach would be more appropriate, i.e. an evaluation based on the independent collection and assessment of appropriate and sound pre-existing supporting evidence. For example, there could be evidence of sufficient volume and quality to permit an evaluation of a method which had long been in use. In other cases, there might not be *in vivo* benchmark data of sufficient breadth and quality to serve as acceptable reference standards for the practical evaluation of an *in vitro* method.

What is weight-of-evidence validation?

The weighing of evidence is, of course, fundamental to the evaluation of any scientific work, but, in this paper, we distinguish between practical validation studies (those involving a multi-laboratory performance of the test) and weight-of-evidence validation studies (those without a dedicated practical study).

This type of validation has been referred to in different ways; for example, *validation by retrospective analysis* (US Interagency Coordinating Committee on the Validation of Alternative Methods, ICCVAM [NIH, 1997]) and *validation through available data* (OECD, 2003). However, we strongly suggest that it would be preferable to call it weight-of-evidence validation, since this more clearly indicates that the approach could be used either retrospectively or prospectively.

There are four principal types of weight-of-evidence validation:
1. The re-evaluation of a previous multi-laboratory validation study (or series of studies).
2. The analysis of data obtained with the same test in different laboratories, but at different times, in studies that were not intended to be parts of a validation exercise.
3. The analysis of data obtained in one or more laboratories using a variation of the protocol that was used in an earlier multi-laboratory validation study.
4. The validation of a testing strategy comprising the use of several test methods, each of which has been previously validated, either in a multi-laboratory validation study or by weight-of-

evidence validation, or a different approach to testing, such as read-across for chemical hazard and risk analysis.

An example of the first type of weight-of-evidence validation would be when a test method is being proposed for a slightly different purpose than that for which it was originally validated. In the second type, it is likely that the protocols used at different times in the various laboratories involved, as well as other protocol parameters, such as sources and types of test chemicals and other materials, would not have been standardised. The third and fourth types of weight-of-evidence validation require judgements to be made about the performances of tests, either when they were combined or when there were small differences in the ways in which they had been conducted.

Criteria for readiness for evaluation and for the quality of the assessment

As with practical validation studies, it must first be established that a procedure or strategy is ready to be evaluated, and, in particular, that there is:

1. a clear definition of the scientific purpose and proposed practical application of the procedure or strategy;
2. a clear description of its scientific basis;
3. a convincing case for its relevance, including an explanation of the need for it in relation to other procedures or strategies;
4. an optimised protocol for the procedure or a clear indication of how the strategy is applied;
5. a statement about the limitations of the approach; and
6. evidence concerning its performance, intra-laboratory reproducibility and, if available, inter-laboratory transferability.

A decision that a procedure or strategy is ready for validation would be taken by a recognised validation authority, such as ECVAM or ICCVAM, who, as sponsors of the study, would appoint a Management Team, who would prepare an overall design for the study. This would proceed in a series of stages.

Stage 1. Appointment of Management Team by Study Sponsors (responsible for overall design of study and appointment of Data Collection and Evidence Assessment groups).

Stage 2. Independent data collection from all available sources by Data Collection Group (to include experienced and independent information scientists and scientists familiar with the type of method and purpose).

Stage 3. Submission of evidence to Evidence Assessment Group and joint review of adequacy of data coverage and quality.

Stage 4. Weight-of-evidence assessment of data by Evidence Assessment Group (according to previously-agreed criteria).

Stage 5. Reporting of outcome to and by Management Team and to Study Sponsors.

Stage 6. Publishing of report of conduct and outcome of study in the peer-review literature.

Stage 7. Independent peer review of study as a whole.

Stage 8. Publication of peer review report.

Stage 9. Consideration by Regulatory Authorities and other appropriate bodies of the acceptability of a validated method for application.

Since the acceptability of the weight-of-evidence assessment itself will also have to be evaluated at a later stage by peer review, it is important that the following criteria are taken into account when the assessment is being planned:

1. Clarity of the defined goals.
2. Quality of the overall design.
3. Independence of management.
4. Quality criteria for evidence.
5. Independence of collection of evidence.
6. Criteria for weighing of evidence.
7. Independence of weighing of evidence procedure.
8. Quality of reporting of outcome.
9. Publication of outcome in peer-reviewed literature.
10. Transparency of whole process (including the identities and affiliations of all the experts involved).

Collecting the evidence

Clearly, the type of evidence to be collected, how it is to be obtained and selected, how its quality is to be checked, and whether it is relevant and reliable, are crucial issues, but it must also be established that the evidence is truly representative of the performance of the procedure or strategy and that its collection is without bias. In addition, how the data are applied, for example, in a prediction model to classify and label chemicals according to a particular type of toxicity, must be included.

It is particularly important that the collection of the evidence is performed by a group of experts who include information scientists and scientists familiar with the type of method under evaluation and the purpose at which it is aimed. This Data Collection Group should be appointed by the Management Team and should be independent both of the developers or proponents of the test procedure or testing strategy and of those who will weigh the evidence once it has been collected.

Weighing the evidence

The performance criteria to be met by a procedure or strategy should be defined by the Management Team in advance of the assessment, according to its proposed purpose, and should be both reasonable and scientifically-based.

The assessment itself cannot be used to improve the evidence but can be used to optimise the use of a method and/or to identify or confirm its strengths and weaknesses. For example, it may become clear that the method is suitable for certain classes of chemicals, but not for others.

It is vital that the members of the Evidence Assessment Group, appointed by the Management Team and responsible for conducting the assessment, have a sufficient breadth and depth of experience and expertise, and that they are independent of the developers or proponents of the method, and of those who collected and presented the evidence.

A case-by-case approach will be essential, and different kinds of evidence will have different levels of value in contributing to

the overall assessment. This will involve evaluations of the plausibility, relevance, consistency, volume and overall strength of the evidence.

Conclusions from a weight-of-evidence assessment

The assessment should lead to a clearly-stated outcome, supported by reasoned argument. There are likely to be three main types of conclusions, depending on the degree to which the weighing of the evidence resolves uncertainty about the relevance and reliability of the test or strategy for its purpose:

1. There is sufficient evidence that a test procedure/testing strategy is reliable and relevant for its stated purpose, and it should be accepted for use for that purpose.
2. There is conflicting evidence about the relevance and reliability of a test procedure/testing strategy for its stated purpose, and a) either the balance is in favour of provisional acceptance or b) further evidence should be obtained.
3. There is sufficient evidence that the test procedure/testing strategy is not reliable and relevant for its stated purpose, and it should not be accepted for use for that purpose.

The outcome of the assessment should be published in a peer-reviewed journal, by or on behalf of the Management Team, as well as being submitted to the sponsors of the exercise and other relevant bodies for an independent and transparent peer review of the study as a whole (design, data collection, weight-of-evidence assessment and reporting).

The application of weight-of-evidence validation

Up to now the European Centre for the Validation of Alternative Methods (ECVAM) has tended to favour a practical approach to validation, whereas the ICCVAM has favoured the weight-of-evidence approach. Thus, of the first ten methods endorsed as valid by the ECVAM Scientific Advisory Committee (ESAC), eight involved practical studies sponsored by ECVAM (tab. 1). The other two endorsements involved methods previously reviewed by ICCVAM.

There is likely to be a trend toward weight-of-evidence assessments, especially as it is increasingly likely that the non-animal tests of the future will contribute evidence that will be used along with other evidence as parts of test batteries and decision-tree testing strategies.

In addition, retrospective validation assessments, based on the ability of tests to give the same predictions as previously obtained, for example with animal tests, will progressively be replaced by prospective assessments, especially where testing strategies are based on more-modern toxicological methods, themselves based on a greater understanding of mechanisms of toxicity and the application of emerging biotechnologies such as toxicogenomics and toxicoproteomics (Bhogal et al., 2005).

Pitfalls to be avoided

It is already clear from experience gained so far that a number of serious pitfalls may be encountered when planning and conducting a weight-of-evidence evaluation, some of which also apply to practical validation studies. These include:

1. implausibility of the test system;
2. inadequate development of the test or testing strategy;
3. lack of evidence and/or poor quality of evidence;
4. bias in selection/presentation of evidence;
5. failure to establish relevance of evidence;
6. lack of a prediction model for applying outcome;
7. lack of clarity/precision in the weighing procedure;
8. pseudosophistication of the weighing procedure;
9. bias in derivation/application of weighing procedure;
10. unreasonably demanding or unreasonably undemanding test performance criteria;
11. injudicious application of the precautionary principle;
12. bias in Data Selection Group or Evidence Assessment Group; and
13. politicisation of the whole process.

The importance and consequences of these pitfalls were discussed at an ECVAM workshop on weight-of-evidence approaches to validation, held in 2004, for which a report is in preparation.

Tab. 1: Endorsement by the ECVAM Scientific Advisory Committee of validated tests for chemical toxicity, 1997-2001.

		Practical study
1. The 3T3 NRU test for phototoxic potential	November 1997	+
2. The EPISKIN™ skin corrosivity test	April 1998	+
3. The rat TER skin corrrosivity test	April 1998	+
4. The application of the 3T3 NRU phototoxicity test to UV filter chemicals	May 1998	+
5. The local lymph node assay for skin sensitisation	March 2000	-
6. The EpiDerm™ skin corrosivity test	March 2000	+
7. The CORROSITEX™ skin corrosivity test	December 2000	-
8. The embryonic stem cell test for embryotoxicity	May 2001	+
9. The whole-embryo culture test for embryotoxicity	May 2001	+
10. The micromass test for embryotoxicity	May 2001	+

Conclusions

There is no doubt that, when permitted to operate effectively and without bias, the ECVAM/ICCVAM/OECD validation process can be used to independently establish that new animal and non-animal test procedures are sufficiently relevant and reliable for their stated purposes and should be considered for regulatory use. However, we have come to the conclusion that the validation process is under threat because of vested interests of various kinds. In addition, it is clear that many currently-accepted animal tests and candidate animal and non-animal tests do not, and could never, meet the agreed criteria for necessity, test development, prevalidation, validation and acceptance.

We have therefore proposed that there is an urgent need for an *invalidation* process to parallel and protect the validation process, so that such methods can be independently reviewed and, where necessary, declared irrelevant and/or unreliable for their claimed purposes (Balls and Combes, 2005). An additional advantage of such a process would be that valuable resources would no longer be wasted in attempts to secure the acceptance of inherently inadequate tests or strategies. This proposal was discussed at an ECVAM/FRAME workshop in September 2005.

Our aim is to ensure that methods will only be put into practice if they have been satisfactorily and independently shown to be reliable and relevant for their stated purposes. We want to protect the reputation of the validation process and, more importantly, prevent the adoption of methods which will ultimately fail, leading to damage to human health.

References

Balls, M., Blaauboer, B., Brusick, D. et al. (1990). Report and recommendations of the CAAT/ERGATT workshop on the validation of toxicity test procedures. *ATLA 18*, 313-337.

Balls, M., Blaauboer, B. J., Fentem, J. H. et al. (1995). Practical aspects of the validation of toxicity test procedures. The report and recommendations of ECVAM workshop 5. *ATLA 23*, 129-147.

Balls, M. and Combes, R. (2005). The need for a formal invalidation process for animal and non-animal tests. *ATLA 33*, 299-308.

Bhogal, N., Grindon, G., Combes, R. and Balls, M. (2005). Toxicity testing: creating a revolution based on new technologies. *Trends Biotechnol. 23*, 299-307.

NIH (1997). Validation and Regulatory Acceptance of Toxicological Test Methods. A Report of the ad hoc Interagency Coordinating Committee on the Validation of Alternative Methods. Research Triangle Park, NC: NIEHS.

OECD (2003). Draft Guidance Document No 34 on the Development, Validation and Regulatory Acceptance of New and Updated Internationally Acceptable Test Methods in Hazard Assessment, Environment Directorate. Paris: OECD.

Correspondence to

Prof. Michael Balls
FRAME
Russell & Burch House
96-98 North Sherwood Street
Nottingham NG1 4EE
UK
e-mail: Michael.Balls@btopenworld.com

Streamlining the Validation Process: The ICCVAM Nomination and Submission Process and Guidelines for New, Revised and Alternative Test Methods*

Leonard M. Schechtman[1], Marilyn L. Wind[2] and William S. Stokes[3]

[1]U.S. FDA, National Center for Toxicological Research, Rockville, Maryland, USA; [2]U.S. Consumer Product Safety Commission, Bethesda, Maryland, USA; [3]NIEHS, Interagency Center for the Evaluation of Alternative Toxicological Methods, Research Triangle Park, North Carolina, USA

Summary

The Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) has developed and implemented a process for the nomination and submission of test methods and for their prioritisation for review and evaluation. Prioritisation of proposed test methods is a function of their regulatory applicability, anticipated multi-agency interest and use, responsiveness to the replacement, reduction, and refinement of animal use, potential for improved predictivity of adverse effects relative to currently employed methods, and efficiency and economic savings. The newly revised ICCVAM Guidelines for the Nomination and Submission of New, Revised, and Alternative Test Methods (http://iccvam.niehs.nih.gov/docs/guidelines/subguide.htm) were developed to assist test method sponsors/nominators in organising the information needed to assess the validation status of test methods at any stage of the validation process and the extent to which the ICCVAM validation and acceptance criteria have been or will be addressed. The original guidelines, in use since 1998 to evaluate the scientific validity of test methods that have since achieved regulatory acceptance, have been updated to reflect experience gained and to help to facilitate a more efficient process. Adherence to these revised guidelines will help ensure the sufficiency of data and information for independent peer review and for regulatory authorities to determine the scientific validity and regulatory acceptability of test methods. The elements comprising these guidelines have now been incorporated into international guidance for the evaluation of methods proposed for new test guidelines. The ICCVAM nomination, submission and prioritisation process and the content and organisation of submissions or nominations are described.

Keywords: ICCVAM guidelines, nominations, submissions, prioritisations

Introduction

The ICCVAM Authorization Act of 2000 stipulates that *Each Federal agency … shall ensure that any new or revised … test method, including animal test methods and alternatives, is determined to be valid for its proposed use prior to requiring, recommending, or encouraging the application of such test method*. In order to facilitate the validation process and provide Federal regulatory agencies with the needed relevant and reliable test methods for regulatory decision-making purposes, ICCVAM was established (initially as an ad hoc committee in 1994, as a standing committee in 1997, and as a permanent body in 2000) and was comprised of 15 Federal regulatory and research agencies. ICCVAM, together with its scientific and operational support centre, NICEATM (the National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods), were created to ensure that new, revised and alternative test methods (a) are scientifically validated to meet agency needs, (b) are more predictive of human health and ecological effects than current methods, and (c) contribute to improved public health. ICCVAM was also tasked with increasing the efficiency and effectiveness of agency test method review, thus eliminating unnecessary duplicative review efforts among different agencies and creating an environment in which agencies could share relevant experiences, information and expertise. Such review efforts would also be aided by optimising the use of scientific expertise recruited from outside the government. ICCVAM was also designated as the principal national body that would promote the replacement, reduction and refinement of animal use for testing and research purposes, the 3Rs (Russel and Burch, 1959).

In order to accomplish these objectives, ICCVAM and NICEATM (a) consider test method nominations and submissions from agencies and the public for review and evaluation of validation status of test methods, (b) review and evaluate new, revised and alternative test methods that may be potentially acceptable for specific regulatory uses, (c) coordinate technical reviews of proposed new or revised or alternative test methods of interagency interest, (d) submit test recommendations to Federal agencies, solicit agency responses, and make these publicly available, and (e) facilitate and/or provide guidance on test method development, validation criteria, validation studies, validation processes, acceptance of scientifically validated test

*The views expressed are solely those of the authors and do not represent any official positions of any Federal agency or organisation.

methods, implementation of ICCVAM-recommended validated methods, awareness of accepted test methods by Federal agencies and other stakeholders, and interagency and international harmonisation of test methods.

Until the advent of ICCVAM, test method validation and acceptance among regulatory agencies had largely been an *ad hoc* procedure, and there was limited communication between agencies regarding validation efforts and processes. ICCVAM responded by devising procedures designed to streamline the validation process and make it more efficient as well as provide a forum for cross-agency communication and consensus-building regarding the evaluation and use of those methods. Although ICCVAM was initially established with a focus on test methods that would address the concept of the 3Rs, it became evident that the validation process and the validation criteria that were to be promulgated by ICCVAM would be equally applicable to all test methods developed for regulatory use, irrespective of whether they were new, revised or alternative in nature.

ICCVAM Guidelines

In 1997, ICCVAM published its Guideline on the *Validation and Regulatory Acceptance of Toxicological Test Methods* (1997). This report provides information on (a) the test method validation process, (b) regulatory acceptance processes, (c) criteria for validation of a test method, (d) criteria for regulatory acceptance of a method, and (e) implementation of a validated test method. The guideline also contains a glossary of terms that would be useful in understanding the terminology used for test method validation purposes and negotiating the validation and regulatory acceptance processes. In 2003, the *ICCVAM Guidelines for Nomination and Submission of New, Revised, and Alternative Test Methods* (2003) were published and were intended to complement the 1997 ICCVAM report. As its title suggests, the document was designed to serve as a tool that would provide guidance to both submitters and regulators on the ICCVAM test method nomination and submission process in order to ensure that nominations/submissions contain adequate information and that the proposed test methods have regulatory applicability. The guideline presents details regarding the ICCVAM criteria for test method prioritisation, validation and regulatory acceptance and it conveys the concept of Performance Standards (Stokes et al., 2006) for test methods. It was developed to guide test method development and to serve as a resource for standardisation and validation efforts. It provides a framework for test method submissions, recommending a skeleton structure outlining the essentials for assisting test method sponsors/nominators in organising the information supporting the validity of a new or modified test method. It describes the information (data, supporting records) needed in test method nominations or submissions to evaluate the validation status of a test method at any stage from development to completion, i.e. the extent to which the validation and acceptance criteria have been or will be addressed.

These two guidelines, together with the encouragement they offer for establishing an ongoing dialogue between sponsors/submitters and ICCVAM/NICEATM, are viewed as a mechanism for facilitating efficient and effective review of submissions and facilitating regulatory acceptance decisions. It is worth noting that the standard outline derived from the ICCVAM nomination and submission guideline is now included in international guidance for submission of test methods proposed for regulatory consideration (OECD, 2005).

ICCVAM nomination, submission and prioritisation process

A test method *nomination* and a test method *submission* differ from each other primarily by the amount of information furnished in support of it. A test method nomination is a test method proposed to ICCVAM for review and evaluation for which a complete test method submission is not available. Examples of nominations are: (a) test methods for which adequate validation studies have apparently been completed but lack a complete submission package; (b) test methods that appear promising based on limited pre-validation or validation data and are proposed for additional validation studies; (c) test methods that have been developed and are proposed for pre-validation or validation studies; and (d) test methods that are recommended for consideration via a workshop or other activity. A test method submission is a test method proposed to ICCVAM for review and evaluation for which adequate validation studies have been completed to characterise the usefulness and limitations of the test method for a specific proposed regulatory testing requirement or application, and adequate documentation of the scientific validity of the test method has been prepared in accordance with ICCVAM test method submission guidelines.

The process followed when ICCVAM receives a nomination to consider the evaluation of a test method or receives a submission of a data package containing complete or partial information needed to assess the validation status of a test method is depicted in figure 1. Upon receipt of a nomination or submission, NICEATM initiates its preliminary evaluation to consider the information received and the extent of the supporting data furnished that would lend credence to the regulatory applicability of the test method. That preliminary evaluation summarises the extent to which proposed test method submissions or nominations address the following ICCVAM prioritisation criteria:

- The extent to which the proposed test method is applicable to regulatory testing needs and applicable to multiple agencies/programmes.
- The degree to which the proposed test method is warranted, based on the extent of expected use or application and impact on human, animal or ecological health.
- The potential for the proposed test method, compared to current test methods accepted by regulatory agencies, to refine animal use (decrease or eliminate pain and distress), reduce animal use, or replace animal use.
- The potential for the proposed test method to provide improved prediction of adverse health or environmental effects, compared to current test methods accepted by regulatory agencies.

- The extent to which the test method provides other advantages (e.g. reduced cost and time to perform) compared to current methods.

NICEATM then presents its recommendations to ICCVAM regarding the next steps that should be considered in the evaluation process. Based upon those recommendations and comments received from its Scientific Advisory Committee on Alternative Toxicological Methods (SACATM) as well as input from the public, ICCVAM determines whether the test method is of sufficient interest and applicability to one or more agencies to warrant further evaluation and renders its recommendations as to the prioritisation that should be assigned to the test method nomination/submission under consideration. NICEATM determines the resources necessary to accomplish ICCVAM's recommendations and management decides whether those resources can be made available for the proposed purpose(s). Upon notification of the availability of the necessary resources, ICCVAM establishes an interagency test method Working Group of knowledgeable scientists to work with NICEATM in organising the necessary evaluation or validation study. Together, ICCVAM, the interagency Working Group, and NICEATM organise workshops, expert panel meetings, independent peer reviews, validation studies, or expedited reviews, as appropriate, to evaluate the validation status of the proposed test method.

Content of a test method submission

The recommended content of a dossier submitted for an ICCVAM assessment of the validation status of a test method is outlined in figure 2. Following is a description of the preferred information that should be included in a test method nomination or submission to ICCVAM. For further, more comprehensive details, the reader is referred to Appendix A of the ICCVAM nomination and submission guideline (ICCVAM, 2003), which provides a detailed outline of the content sought for nominations and submissions to ICCVAM. Although there is no mandatory minimum requirement regarding the information to be provided, ICCVAM's consideration of the proposed test method will be expedited by providing as much information as possible. Furthermore, experience has demonstrated that the more closely a submission adheres to reporting the suggested information in a format approximating that of the suggested outline, the more efficient is the evaluation process and the fewer the data gaps that would be encountered and that would need filling after the fact. Areas where the requested information is not available or is incomplete should be indicated and the studies that would be conducted to generate those necessary data should be specified.

Section 1 provides the introduction and rationale for the proposed test method. It introduces the proposed test method and describes its regulatory and scientific rationale. This would include a description of how the proposed test method can be used in a regulatory setting (e.g. as a screen in a tiered testing strategy, as an adjunct test to provide mechanistic information, as a surrogate for an existing test method) and how the proposed test method might be included in the overall safety or hazard assessment process. This information would also include a discussion of the mechanistic basis of the proposed test method and the context in which it will be used to measure or predict the toxicological activity of a test material or substance.

Section 2 reports on the test method protocol components. This would include an explanation and description of the basis for decisions on critical functional, structural, and procedural ele-

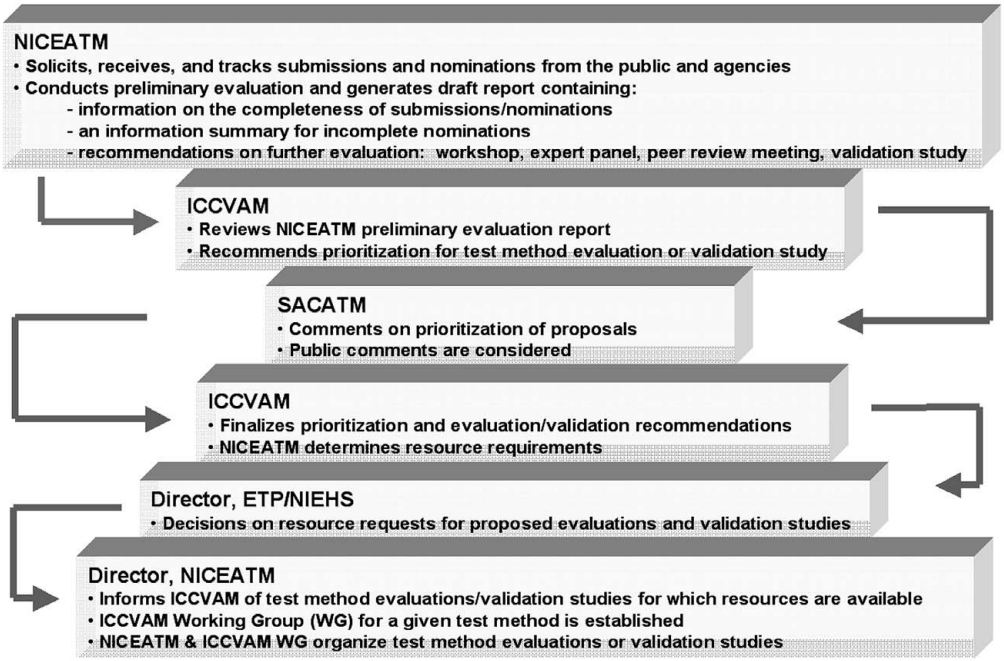


Fig. 1: ICCVAM test method nomination, submission, and prioritisation process

ments of the test method protocol. Critical protocol elements would include such factors as (a) dose/concentration selection procedures, (b) duration of exposure, (c) number of replicate and/or repeat experiments, (d) endpoint(s) measured, (e) concurrent (positive, negative, vehicle) controls and the basis for their selection, (f) acceptable response range, (g) statistical or other methods of data analysis (including dose-response relationships if applicable), (h) decision criteria, i.e. those criteria used to designate a response as positive, negative or equivocal, (i) test system specifications, e.g. cell systems and properties, animal models (species, strain, sex, age, genetic background, diet), (j) details regarding any proprietary components of the test method and the maintenance of their integrity (reliability and accuracy). In addition, a complete, detailed protocol for the proposed test method should be appended to the submission.

Section 3 addresses the Substances Used for Validation of the Proposed Test Method. Here, the rationale for the numbers and types of substances tested during the validation process and the specific chemical or formulation names and relevant chemical and product classes are reported. It is important to recognise that the limitations associated with a test method are often indicative of its effectiveness in evaluating certain classes of chemicals and therefore any characteristics thought to have direct impact on test method accuracy and/or reliability should be described. To the extent possible, the information provided for each test substance should include (a) Chemical Abstracts Service Registry Number (CASRN), (b) physical and chemical characteristics, (c) concentrations tested, (d) purity, (e) source, (f) stability of the test substance in the test medium, and (g) in the case of mixtures, the constituents and their relative concentrations. Information regarding the use of coded substances and blind testing during the validation process should also be reported. In cases where the proposed test method is mechanistically and functionally similar to a previously validated ("pioneer") test method, a discussion should be presented regarding the performance of the proposed ("me-too") test method using reference chemicals recommended in the Performance Standards established for the pioneer test method.

1.0 Introduction and scientific and regulatory rationale for the proposed test method
2.0 Proposed test method protocol
3.0 Substances used for validation
4.0 Reference data used for performance assessment
5.0 Test method data and results
6.0 Test method accuracy assessment
7.0 Test method reliability (repeatability/reproducibility) assessment
8.0 Test method data quality
9.0 Other scientific reports, reviews pertinent to the proposed test method
10.0 Animal welfare considerations: refinement, reduction, and replacement
11.0 Practical considerations (cost, time, training, equipment, transferability)
12.0 References
13.0 Supporting materials (e.g., detailed test method protocol) in appendices

Fig. 2: Content of a test method submission

Section 4 discusses the *in vivo* reference data used for the performance assessment of the proposed test method. If the proposed test method is intended to replace or substitute for an existing *in vivo* reference test method, then a comparison of data generated by each is essential. Information sought would also include such factors as (a) the protocol(s) used to generate the *in vivo* reference data, (b) the *in vivo* reference data (human, animal) used to assess the accuracy of the proposed (new, modified) test method, (c) the quality of the *in vivo* reference test method data and the extent of Good Laboratory Practice (GLP) compliance (U.S. EPA, 2003a; U.S. EPA, 2003b; U.S. FDA, 2003; OECD, 1998; Japanese Good Laboratory Practice Standards, 1997; Cooper-Hannan et al., 1999), (d) the criteria used to select the *in vivo* reference test method (or human) data, (e) the availability of original study data for the *in vivo* reference test method studies reported.

Section 5 conveys the test method data and results. Here, the data generated by the testing of chemicals and substances using the proposed test method protocol are reported, including all data (accompanied by "raw" data) from all studies (successes and failures), summary data presented in tabular or graphic form as appropriate, the statistical approach(es) used, together with their justification, in the analysis of the data. Accompanying those data would be discussions of protocol modifications, if any, that might have been introduced during method development and their impact.

Section 6 presents the assessment of the performance of the proposed test method relative to the currently accepted reference method. The accuracy of the test method (e.g. its concordance, sensitivity, specificity, false positive and false negative rates, positive and negative predictivity) is calculated and discussed with respect to its ability to measure or predict the effect of interest and any results that may be discordant with those generated by the reference test method would be discussed. In instances where the proposed test method is measuring or predicting an endpoint for which there is no pre-existing test method, the frequency of correct predictions should be compared to relevant information from the species of interest. In cases where the proposed test method is a me-too method, i.e. is mechanistically and functionally similar to a validated (pioneer) test method with established performance standards, the accuracy of both test methods should be compared. Also addressed in this section would be a discussion of the usefulness and limitations of the test method based on its performance characteristics.

Section 7 provides an assessment of the reliability (repeatability and reproducibility) of the test method. Here, analyses and conclusions would be presented for intra-laboratory repeatability and intra- and inter-laboratory reproducibility together with a quantitative statistical analysis of the extent of any variability encountered. This assessment would include a discussion of the rationale for the selection of the substances used to evaluate the intra- and inter-laboratory reproducibility and the extent to which those substances represent the range of possible test outcomes. Historical negative, vehicle and positive control data and

the responses of the proposed test method relative to these control values would also be presented and discussed. In cases where the proposed test method is a me-too method similar to a validated pioneer test method with established performance standards, the reliability of both test methods should be compared and the potential impact of any differences discussed.

Section 8, which discusses test method data quality, communicates the extent of adherence to national and international GLP guidelines (OECD, 2005; U.S. EPA, 2003; U.S. EPA, 2003b; U.S. FDA, 2003; OECD, 1998; Japanese Good Laboratory Practice Standards, 1997) as well as the results of any data quality audits. Deviations from GLP guidelines and their impact, unpublished data (supported by laboratory notebook entries), and the availability of laboratory notebooks and other retained data would also be addressed.

Section 9, which deals with other relevant scientific reports and reviews, discusses data and reports from other studies (published and unpublished) conducted using the proposed test method. Conclusions regarding the proposed test method available from independent peer-reviewed reports or other scientific reviews would be presented and those conclusions should be compared to those reached in the submission. In cases where the proposed test method is a me-too method similar to a validated pioneer test method with established performance standards, the results of any studies conducted subsequent to the ICCVAM evaluation should be reported and any impact on the reliability and accuracy of the proposed test method would be discussed.

Section 10 addresses animal welfare considerations. In this section a description of how the method will refine, reduce or replace animal use compared to currently accepted methods used for the endpoint of interest is presented. If animals are used in the test method, the discussion addresses (a) the rationale for the use of animals, (b) the basis for the number of animals required, (c) approaches used to prevent or minimise pain and distress as appropriate and (d) sources used to determine availability of 3Rs alternatives (USDA, 1966; Public Health Service (1996).

Section 11, which addresses practical considerations, includes discussions of the transferability aspects of the proposed test method (e.g. facilities, equipment, supplies, expertise, training and proficiency necessary), as well as time and cost considerations.

Section 12, references, cites all publications referenced in the submission.

Section 13, supporting materials, is comprised of all necessary appendices containing information relevant to the proposed test method that would supplement the central content of the submission. Such material would include, but would not necessarily be limited to, (a) the complete detailed protocol and relevant standard operating procedures (SOPs) as appropriate, (b) copies of all relevant publications, (c) all non-transformed original data, (d) suggested performance standards developed for the proposed test method, and (e) any other supporting documentation considered informative and helpful.

Conclusion

The *ICCVAM Guidelines for the Nomination and Submission of New, Revised, and Alternative Test Methods* (ICCVAM, 2003), were developed to assist test method sponsors and nominators in organising the information needed by ICCVAM to assess the validation status of a new or modified test method at any stage of the validation process. These guidelines include guidance on the process for submitting nominations to ICCVAM for test methods that are proposed for further consideration, but which may require further compilation of data or even additional validation studies. The guidelines convey the ICCVAM test method nomination and submission process. They also describe the information that should be provided in test method nominations or submissions and provide a framework for that information so that ICCVAM can evaluate appropriately the extent to which the validation and acceptance criteria have been addressed or will be addressed in proposed studies. In addition, the guidelines put into practice the use of performance standards, which communicate the basis on which a validated and accepted proprietary (i.e. copyrighted, trademarked, registered) or non-proprietary test method has been determined to have sufficient accuracy and reliability for a specific testing purpose. Such performance standards are designed to be met by proposed test methods that are based on similar scientific principles and that measure or predict the same biological or toxic effect.

Test method sponsors are encouraged to employ these guidelines together with the ICCVAM guideline on the *Validation and Regulatory Acceptance of Toxicological Test Methods* (ICCVAM, 1997) and to seek consultative advice from ICCVAM and NICEATM throughout the test method development, prevalidation, and validation process, as well as during preparation of submissions. This interactive process enhances the likelihood that agencies will have sufficient data and information to determine the extent that a test method can generate information that will meet their regulatory needs. Proper usage of these guidelines and capitalisation on interactions with ICCVAM and NICEATM will help to maximise the likelihood that validation studies and submissions will adequately characterise the usefulness and limitations of the proposed test method. Complete submissions are essential and serve as a basis for assessing the validation status of a proposed test method through an independent ICCVAM peer review process.

References

Cooper-Hannan R, Harbell J, Coecke S, et al. (1999). The principles of good laboratory practice: application to in vitro toxicology studies. *ATLA 27*, 539-577.

ICCVAM Authorization Act of 2000. (2000). Public Law 106-545, December 19, 2000. Available: http://iccvam.niehs.nih.gov/about/PL106545.pdf [accessed 1 November 2005].

ICCVAM. (1997). Validation and Regulatory Acceptance of

Toxicological Test Methods: A Report of the ad hoc Interagency Coordinating Committee on the Validation of Alternative Methods. NIH Publication No. 97-3981. National Institute of Environmental Health Sciences, Research Triangle Park, NC. Available: http://iccvam.niehs.nih.gov/docs/docs.htm#general [accessed 1 November 2005].

ICCVAM. (2003). ICCVAM Guidelines for the Nomination and Submission of New, Revised, and Alternative Test Methods. NIH Publication No. 03-4508. National Institute of Environmental Health Sciences, Research Triangle Park, NC. Available: http://iccvam.niehs.nih.gov/docs/guidelines/subguide.htm [accessed 1 November 2005].

Japanese Good Laboratory Practice Standards, including: (1) Pharmaceutical GLP Standard. (1997). Pharmaceutical Affairs Bureau, Ministry of Health and Welfare, Available: http://www.jsqa.com/en/glp/glp21e.html [accessed 1 November 2005]; and (2) Good Laboratory Practice Standards. 1984. Japanese Ministry of Agriculture, Forestry, and Fisheries. Available: http://www.ovpr.uga.edu/qau/maffintr.html [accessed 1 November 2005].

OECD (1998). Principles of Good Laboratory Practice and Compliance Monitoring. ENV/MC/CHEM (98)17. Paris, France: Organisation for Economic Co-operation and Development. Available: http://www.olis.oecd.org/olis/1998doc.nsf/LinkTo/env-mc-chem(98)17 [accessed 1 November 2005]. Also available at: http://www.oecd.org/document/63/0,2340,en_2649_34381_2346175_1_1_1_1,00.html [accessed 1 November 2005].

OECD. (2005). OECD Series on Testing and Assessment, Number 34. Guidance Document on the Validation and International Acceptance of New or Updated Test Methods for Hazard Assessment, Paris, France. Available: http://appli1.oecd.org/olis/2005doc.nsf/linkto/env-jm-mono(2005)14 [accessed 1 November 2005].

Public Health Service. (1996). Public Health Service Policy on Humane Care and use of Laboratory Animals. Washington, DC, U.S. Department of Health and Human Services. Available: http://grants1.nih.gov/grants/olaw/references/phspol.htm [accessed 1 November 2005].

Russell, W. M. S. and Burch, R. L. (1959). The Principles of Humane Experimental Technique. London: Methuen.

Stokes, W. S., Schechtman, L. M., Wind, M. et al. (2006). The Use of Test Method Performance Standards to Streamline the Validation Process, *ALTEX*, this issue.

U.S. EPA. (2003a). Federal Insecticide, Fungicide, and Rodenticide Act: Good Laboratory Practice Standards; Final Rule. 40 CFR Part 160. Available: http://www.access.gpo.gov/nara/cfr/waisidx_03/40cfr160_03.html [accessed 1 November 2005].

U.S. EPA. (2003b). Toxic Substances Control Act (TSCA): Good Laboratory Practice Standards; Final Rule. 40 CFR Part 792. Available: http://www.access.gpo.gov/nara/cfr/waisidx_03/40cfr792_03.html [accessed 1 November 2005].

U.S. FDA. (2003). Good Laboratory Practice for Nonclinical Laboratory Studies. 21 CFR Part 58. Available: http://www.access.gpo.gov/nara/cfr/waisidx_03/21cfr58_03.html [accessed 1 November 2005].

USDA. (1966). Animal Welfare Act. 7 USC: 2131-2156. Available: http://warp.nal.usda.gov:80/awic/legislat/usdaleg1.htm [accessed 1 November 2005].

Correspondence to
Leonard M. Schechtman, Ph.D.
U.S. FDA
National Center for Toxicological Research
Rockville
Maryland, USA
e-mail: leonard.schechtman@fda.hhs.gov

The Use of Test Method Performance Standards to Streamline the Validation Process

William S. Stokes[1], Leonard M. Schechtman[2], Amy Rispin[3], Kailash Gupta[4], Karen Hamernik[3], Raymond Tice[1] and Marilyn Wind[4]

[1]National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods, National Institute of Environmental Health Sciences, NIH, DHHS, Research Triangle Park, NC, USA; [2]US Food and Drug Administration, National Center for Toxicological Research, Rockville, MD, USA; [3]U. S. Environmental Protection Agency, Washington, DC, USA; [4]Consumer Product Safety Commission, Bethesda, MD, USA

Summary

The Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) recently defined and established a process for the development and use of test method performance standards. Performance standards communicate the basis by which proprietary and non-proprietary test methods have been determined to be acceptable for proposed specific testing uses. They also provide the basis for evaluating the acceptability of proposed test methods that are structurally and functionally similar to an accepted test method. The development and use of test method performance standards is expected to streamline the validation and regulatory acceptance of new, revised and alternative test methods.

Keywords: performance standards, validation, regulatory acceptance, proprietary test methods

Introduction

Alternative test methods proposed for regulatory use must undergo validation studies to assess their reliability and relevance for specific applications (ICCVAM, 1997, 2003; OECD, 2005a). Regulatory agencies use this information to determine the applicability and acceptability of the test method for addressing specific regulatory safety testing requirements. Many new and alternative test methods are proprietary in nature and are protected by intellectual property laws such as patents, trademarks, and copyrights. Such intellectual protections stimulate innovation by providing financial incentives for companies to develop and market new products, such as *in vitro* testing methods that may reduce, refine, or replace animal use. However, US laws require that government regulatory authorities cannot simply endorse or approve proprietary methods until they first convey the basis by which the proprietary methods have been determined to be acceptable for use (CFR, 2005; EPA, 1995).

The requirement for regulatory agencies to provide the basis for acceptability of a proprietary test method was first brought to the attention of the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) during USEPA (US Environmental Protection Agency) consideration of four *in vitro* test methods for skin corrosion, of which three were proprietary. All four *in vitro* test methods had been reviewed and recommended by ICCVAM for identifying chemicals that may cause skin corrosivity (ICCVAM, 1999, 2002). The USEPA requested that ICCVAM develop performance standards that could be used by the USEPA to convey the basis by which each of these four *in vitro* methods were found to be useful as a screening test to identify dermal corrosives in a tiered testing strategy.

At about the same time, ICCVAM proposed a generic version of Corrositex® as an OECD test guideline. This generic version was necessary since the OECD Test Guidelines Programme does not allow inclusion of any proprietary test methods or test method components in OECD test guidelines. However, it was not clear how OECD member countries that received data from a generic version of Corrositex® would be able to determine that the generic version was sufficiently valid such that it performed as well as or better than the proprietary test method on which the generic test guideline was based.

In order to address the issues relevant to both the regulatory agency needs and the OECD test guideline issues, an ICCVAM initiative was undertaken to develop and define the concept of test method performance standards that could be used for both situations. This paper will discuss the critical elements of performance standards and present the process by which ICCVAM develops and validates performance standards. An example of successful performance standards will also be presented.

Defining test method performance standards

ICCVAM defined performance standards as the basis on which a proprietary or non-proprietary test method has been determined to have sufficient accuracy and reliability for a specific testing purpose (ICCVAM, 2003). This process involves determination that a test method has sufficient accuracy and reliability for a defined specific testing purpose, and then using this information to develop performance standards that can be used

as the basis for evaluating the acceptability of proposed test methods that are based on similar scientific principles and that measure or predict the same biological or toxic effect. These performance standards can then be used by regulatory authorities to communicate the basis by which they find test methods to be acceptable for specific regulatory testing purposes. ICCVAM now routinely develops and proposes performance standards during test method evaluations for both proprietary and non-proprietary methods that have undergone adequate validation (ICCVAM, 2003).

Components of performance standards

Performance standards have three critical components: essential test method components, a minimum list of reference chemicals, and defined accuracy and reliability values (ICCVAM, 2003). If a similar test method adequately addresses and meets these standards, then it would be considered to be comparable, in terms of performance, to the test method used to establish the performance standards.

Essential test method components

are the requisite structural, functional, and procedural elements of a validated test method that should be included in the protocol of a proposed mechanistically and functionally similar test method. These components include unique characteristics of the test method, critical procedural details, and quality control measures. If there are deviations from the recommended essential test method components, then a scientific rationale must be provided and any potential impact of the deviations must be discussed. Incorporation of and adherence to essential test method components will help assure that a proposed test method is based on the same concepts as the corresponding validated test method.

The minimum list of reference chemicals

is used to assess the accuracy and reliability of a mechanistically and functionally similar test method that incorporates all of the essential test method components. These chemicals are a representative subset of those used to demonstrate the reliability and accuracy of the validated test method. To the extent possible, these reference chemicals should:

- Represent the range of responses that the validated test method is capable of measuring or predicting (e.g. negative and weak to moderate to strong positives)
- Produce consistent results in the validated test method and in the *in vivo* reference test method and/or target species of interest
- Reflect the accuracy of the validated test method
- Have well-defined chemical structures
- Be readily available, i.e. can be purchased from commercial sources
- Not be associated with excessive hazard or prohibitive disposal costs
- Represent the range of known or suspected mechanisms and/or modes of action for the toxicity measured or predicted by the test method
- Represent the range of physical and chemical properties which the test method is proposed to be capable of testing (e.g. solubility, pH, volatility, etc.)

These reference chemicals are the minimum number that should be used to evaluate the performance of a proposed mechanistically and functionally similar test method. These chemicals should not be used to develop the decision criteria/prediction model for the proposed test method. If any of the recommended chemicals are unavailable, other chemicals for which adequate reference data are available could be substituted with adequate scientific justification. To the extent possible, any substituted chemical(s) should be of the same chemical class and potency as the original chemical(s). If desired, additional chemicals representing other chemical or product classes and for which adequate reference data are available can be used to more comprehensively evaluate the accuracy of the proposed test method. However, these additional chemicals should not include those used to develop the proposed test method.

Accuracy and reliability values

are the comparable performance that should be achieved by the proposed test method when evaluated using the minimum list of reference chemicals. Reference chemicals should be designated for performance standards that will result in accuracy and reliability values similar to the overall values determined from the entire validation database for the reference test method.

Process for developing performance standards

ICCVAM has developed a process for establishing performance standards during the evaluation of proposed new test methods (ICCVAM, 2003). The process is designed to ensure rigorous scientific review and to provide the opportunity for broad stakeholder and public comment. The ICCVAM process for developing performance standards for new test methods is as follows:

- The National Toxicology Program Interagency Center for the Evaluation of Alternative Methods (NICEATM) and the appropriate ICCVAM working group develop proposed performance standards for consideration during the ICCVAM evaluation process. If a sponsor proposes performance standards, these are considered by ICCVAM at this stage. Generally, the proposed performance standards will be based on the information and data provided in the test method submission or on other available applicable data.
- The ICCVAM/NICEATM Peer Review Panel evaluates the proposed performance standards for completeness and appropriateness during its evaluation of the validation status of the proposed test method. The proposed performance standards are made available with the test method submission to the public for comment prior to and during the Peer Review Panel meeting.
- The appropriate ICCVAM working group, with the assistance of NICEATM, prepares the final performance standards for ICCVAM approval, taking into consideration the recommendations of the Peer Review Panel and public comments.
- Performance standards recommended by ICCVAM are incor-

porated into ICCVAM test method evaluation reports, which are published, provided to Federal agencies, and made available to the public. Availability of ICCVAM test method evaluation reports are announced routinely in the *Federal Register*, NTP Newsletters, and ICCVAM/NICEATM e-mail list serve groups.

- Regulatory authorities can then reference the performance standards in the ICCVAM report when they communicate their acceptance of a new test method. In addition, performance standards adopted by regulatory authorities can be provided in guidelines issued for new test methods.

Performance standards for dermal corrosivity test methods

Following the concept development and definition of performance standards, ICCVAM then used this framework to address the request by the USEPA to develop performance standards for the three proprietary dermal corrosivity test methods previously recommended by ICCVAM: Corrositex®, Episkin™, and Epiderm™ (ICCVAM 1999, 2002). In addition, and in accordance with the adopted process for developing performance standards, ICCVAM also developed performance standards for the one recommended non-proprietary dermal corrosivity test method, the rat skin transcutaneous electrical resistance (TER) method.

Due to the structural and functional differences of the four methods, three different sets of performance standards were developed (ICCVAM, 2004). Since Episkin™, and Epiderm™ are structurally and functionally similar, one set of performance standards was developed for these two methods. The standards were based on Episkin™, since this method had a larger validation database than Epiderm™ (60 vs. 24). In addition to the essential test method components, a minimum list of 24 reference chemicals were selected from the 60 chemicals used for the validation of Episkin™. This included 12 corrosives and 12 non-corrosives. All of the selected reference chemicals are commercially available. Accuracy and reliability values for the 24 minimum reference chemicals closely matched the overall performance for the 60 chemicals in the validation database. For the rat skin TER, a minimum list of 24 reference chemicals was selected, which also provided similar accuracy and reliability values as those for the total validation database of 60 chemicals.

Performance standards based on Corrositex® were developed for a generic *in vitro* membrane barrier test system for skin corrosion (ICCVAM, 2004). The essential test method components were included in the test method description for an OECD test guideline based on the validated Corrositex® (OECD, 2005b). This test method is capable of identifying the three subcategories of corrosivity described by the United Nations Packing Group (PG) classification system. Accordingly, the validation database contained a larger number of substances (129). The selected minimum list of reference chemicals contained a total of 40 chemicals, including 12 non-corrosive methods and 28 corrosive chemicals (9 UN PGI, 9 UN PGII and 10 UN PG III). As with the other methods, the accuracy and reliability values for the minimum list of reference chemicals were similar to those for the total validation database. These performance standards were subsequently included in the proposed OECD Test Guideline 435 for an *in vitro* membrane barrier test system for skin corrosion (OECD, 2005b). This is the first OECD test guideline to incorporate all three critical performance standard components.

Using performance standards for validation studies

The availability of performance standards is expected to significantly expedite the validation and acceptance of improved test methods that are similar to previously accepted methods.

Since validation can be a lengthy and expensive process, it is clearly advantageous to reduce the time and expense required to determine the usefulness and limitations of new test methods, especially those that may be structurally and functionally similar to test methods for which there has already been adequate validation. In fact, the idea that performance criteria could be used for this purpose has been previously proposed (Balls, 1997). Obviously the complete and full validation of structurally and functionally similar test methods could be expensive and potentially unnecessarily duplicative. For example, without performance standards, evaluation studies on a generic version of Corrositex®, which underwent validation with over 129 chemicals (ICCVAM, 1999), might be expected to use a similar number of chemicals. The establishment of performance standards that can be used to determine if similar test methods have comparable or better performance than one that has undergone extensive validation and been determined to be acceptable for a specific regulatory testing purpose will save considerable time and expense. The availability of performance standards should also facilitate the validation of improved versions of tests that may provide for better predictivity and efficiency. The concept and definition of test method performance standards has also recently been included in international guidance on validation (OECD, 2005a).

Conclusions

The establishment of test method performance standards now permits the regulatory acceptance of proprietary test methods and test methods with proprietary components. Test method performance standards have now been successfully developed for three different types of *in vitro* dermal corrosivity test methods. Performance standards have also now been incorporated into international test guidelines and international guidance on validation and regulatory acceptance. The continued development and availability of test method performance standards can be expected to streamline both the validation and acceptance processes for improved test methods that are mechanistically and functionally similar to previously accepted methods.

References

Balls, M. (1997). Editorial: Defined Structural and Performance Criteria would Facilitate the Validation and Acceptance of Alternative Test Procedures. *ATLA 25*, 483-484.

Environmental Protection Agency (1995). Ethics Advisory 95-18, Use of Public Office for Private Gain – Endorsements. Based on provisions of the "Standards of Ethical Conduct for Employees of the Executive Branch" at 5 CFR 2635.702 and on provisions of 5 CFR 2635.703. Washington, DC, USA: US Environmental Protection Agency.

ICCVAM (1997). Validation and Regulatory Acceptance of Toxicological Test Methods: A Report of the ad hoc Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM), NIH Publication No. 97-3981, NIEHS, Research Triangle Park, NC, USA. (URL: http://iccvam.niehs.nih.gov/docs/guidelines/validate.pdf).

ICCVAM (1999). Corrositex®: An In Vitro Test Method for Assessing Dermal Corrosivity Potential of Chemicals, NIH Publication No. 99-4495, NIEHS, Research Triangle Park, NC, USA. (URL: http://iccvam.niehs.nih.gov/docs/reports/corprrep.pdf).

ICCVAM (2002). ICCVAM Evaluation of EPISKIN™, EpiDerm™, (EPI-200), and Rat Skin Transcutaneous Electrical Resistance (TER): In Vitro Test Methods for Assessing Dermal Corrosivity Potential of Chemicals. NIH Publication No. 02-4502. NIEHS, Research Triangle Park, NC, USA. (URL: http://iccvam.niehs.nih.gov/methods/epiderm.htm).

ICCVAM (2003). ICCVAM Guidelines for the Nomination and Submission of New, Revised, and Alternative Test Methods. NIH Publication No. 03-4508. NIEHS, Research Triangle Park, NC USA. (URL: http://iccvam.niehs.nih.gov/docs/guidelines/subguide.htm).

ICCVAM (2004). ICCVAM Recommended Performance Standards for In Vitro Test Methods for Skin Corrosion. NIH Publication No. 04-4510. NIEHS, Research Triangle Park, NC USA May 2004. (URL: http://iccvam.niehs.nih.gov/docs/docs.htm#invitro).

OECD (2005a). Guidance Document No. 34: Guidance Document on the Validation and International Acceptance of New or Updated Test Methods for Hazard Assessment. Paris, France: Organization for Economic Cooperation and Development.

OECD (2005b). In Vitro Membrane Barrier Test Method for Skin Corrosion, Test Guideline 435. Organization for Economic Cooperation and Development, Paris, France. (URL: http://www.oecd.org).

U.S. Code of Federal Regulations (2005). "Standards of Ethical Conduct for Employees of the Executive Branch" at 5 CFR 2635.702. Washington, DC, USA: US Government Printing Office.

Acknowledgements

The authors wish to acknowledge the many designated agency representatives from the 15 member agencies of ICCVAM who contributed to the development and review of the ICCVAM guidance on performance standards, and members of the ICCVAM Dermal Corrosivity and Irritation Working Group who contributed to development of performance standards for *in vitro* methods for the *in vitro* corrosivity methods. Finally, the authors wish to acknowledge the expert assistance of Ms. Debbie McCarley in the preparation of this manuscript.

Correspondence to

Dr. William S. Stokes
National Institute of Environmental Health Sciences
Mail Drop EC-17, Post Office Box 12233
79 T. W. Alexander Drive
Research Triangle Park, North Carolina
USA 27709
e-mail: stokes@niehs.nih.gov

Workshop 5.14
Meeting the challenge of the 7th amendment to the EU cosmetics directive (COLIPA Workshop)

Good Science Must be the Key Factor in the Development and Use of Alternative Methods for Safety Assessment of Cosmetics

Vera Rogiers and Marleen Pauwels
Dept. of Toxicology, Vrije Universiteit Brussel, Belgium

Summary
The most striking features of the current EU cosmetic legislation are its animal testing and marketing ban. As the legally imposed timeframe has mainly been introduced by political considerations, the key question remains whether replacement of all in vivo tests is scientifically feasible.
The overview of the current status of 3R alternatives and related deadlines provided by ECVAM were commented by the SCCNFP.
In this paper, repeated dose toxicity testing is discussed to show that innovation is taking place but cannot be governed by political pressure. The opening up of the field of alternatives and the creation of a positive atmosphere towards good science are highly needed.

Keywords: science, alternative methods, safety assessment, cosmetics

Introduction

Both the 7th Amendment of the EU Cosmetic Directive 2003/15/EC (OJ L66, 11/03/2003) and the REACH (Registration, Evaluation and Authorisation of Chemicals) proposal for chemicals (Anon., 2003a) assume that alternatives to animal tests for regulatory purposes are or will be available within the near future.

In particular, the 7th Amendment imposes a prohibition, not only on testing finished cosmetic products and cosmetic ingredients on animals, but also on marketing such products in the EU.

The testing ban on finished products applies from 11 September 2004. So does the testing ban on ingredients, though only when alternative methods become available, i.e. tests that have been officially validated by ECVAM. The maximum cut-off date for the testing ban is 11 March 2009, irrespective of the availability of non-animal alternatives. The marketing ban will be introduced at the latest on the same date, but with the exception of repeated dose toxicity, reproductive toxicity and toxicokinetic testing. For those, the deadline is fixed on 11 March 2013. In the REACH proposal (Anon., 2003a), it is indicated that a scientific objective for the EU is the development and validation of alternative methods and that such methods must be considered as and when they become available. Thus an important difference between both is that in the chemical legislation proposal the 3Rs concept of Russell and Burch (1992) is fully considered, whereas in the cosmetic legislation only replacement was taken up.

The concept of using alternative methods to animal testing was first introduced in the EU legislation by Directive 86/609/EEC (OJ L358, 18/12/1986) on the protection of animals used for experimental and other purposes, which was the immediate reason why ECVAM (the European Centre for the Validation of Alternative Methods) was established in 1992. Alternatives were meant to include all of the 3Rs, being methods to replace, reduce or refine animal tests. Only in the cosmetic legislation was this inscribed and subsequently turned into 1R as a direct consequence of political pressure and excessive lobbying of the Parliament. Thus it is important to notice that the actual deadlines of 2009 and

2013 were not introduced owing to new scientific developments or break-through methodologies becoming available, but rather on account of political considerations.

The crucial question therefore remains whether development of all alternatives needed to guarantee safety for human health is scientifically spoken possible within the limited timeframe foreseen.

Safety evaluation of cosmetics according to the EU cosmetic legislation

In the EU, cosmetics must be safe for human health and the responsibility for this lies with the manufacturer, marketer or first importer into the EU market. The safety assessment of cosmetics takes into account the chemical structure of all ingredients, their toxicological profile and exposure. Toxicological testing of cosmetic ingredients/chemicals is traditionally done using experimental animals, but nowadays preference is given, whenever possible, to alternative methods. Appropriate safety data must be available for all ingredients, regardless of the tonnage of marketing within the EU, in order to permit safety evaluation (= risk assessment) of the finished cosmetic product (Pauwels and Rogiers, 2004).

In Europe, two distinct channels are functional (fig. 1) (Anon., 2003b). Basically, the substances present in Annexes II, III, IV, VI and VII of the Cosmetic Directive fall under the responsibility of the SCCP (Scientific Committee on Consumer Products, the former SCCNFP) (left part of fig. 1).

The right part of figure 1 contains all ingredients of cosmetic products other than those present in the Annexes. They fall under the responsibility of the manufacturer through the safety assessor. In fact, a specific data package must be kept readily accessible to the EU Member States' Competent Authorities. This is called a Technical Information File (TIF) or a Product Information Requirement (PIR) (Pauwels and Rogiers, 2004).

The set of tests to be considered for the determination of the toxic potential of the ingredients present in the Annexes is represented in table 1 (Anon., 2003b).

For the ingredients not present in the annexes, usually acute toxicity, local toxicity, dermal absorption, repeated dose toxicity and mutagenicity are considered to be the minimal base set requirements (Anon., 2003b). According to Directive 92/32/EEC (OJ L54, 05/06/1192) on the classification, packaging and labelling of dangerous substances, the toxicological requirements for substances produced/EU imported at levels between 100 kg and 1 tonne per year consist of the minimum set

Tab. 1: General toxicological requirements for cosmetic ingredients, present in the Annexes of the Cosmetic Directive.

Acute toxicity (if available)
Irritation and corrosivity
Skin sensitisation
Dermal absorption
Repeated dose toxicity
Mutagenicity/genotoxicity
Carcinogenicity
Reproductive toxicity
Toxicokinetics
Photo-induced toxicity
Human data

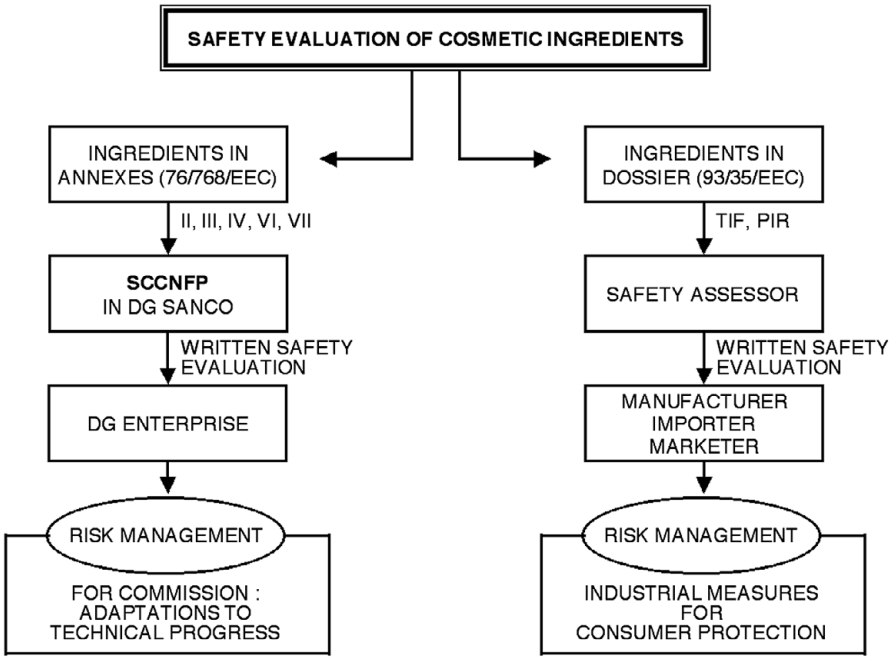


Fig. 1: Two distinct channels for safety assessment of cosmetic ingredients are functioning in the EU (Anonymous, 2003b).

just mentioned, but without dermal absorption and repeated dose toxicity. Thus, not all of the test results required for the assessment of cosmetic ingredients are automatically generated through the chemicals' legislative procedures. Only when higher amounts of substances are involved, a more extensive list of toxicological requirements is established.

Alternatives to animal testing

Alternatives do not yet exist for all toxicological tests required to guarantee the safety of cosmetic ingredients for human health (Anon., 2003b). A recent overview of the current status and future prospects of alternatives has been prepared by ECVAM (Anon., 2004a) and was recently published (Eskes and Zuang, 2005). It provides a good overview of all 3R alternatives that exist or are under development.

The original ECVAM document was commented upon by the former SCCNFP (Anon., 2004b) and the following clear message was given to the Commission:

-"total abolishment of animal tests within 10 years is not feasible from an objective scientific point of view. Even the alternative strategies discussed in the document, which are estimated to take more than 10 years for further development, still include an animal test in the finalities."

-"In many cases (especially for systemic toxicity), all the gaps are not identified."

-"Following fields are missing throughout the document:

- absorption through mucosa
- respiratory sensitisation
- in the repeated dose toxicity and developmental toxicity sections, two of the "3R" (refinement and reduction) have been neglected
- upcoming alternative technologies, such as the use of stem cells, genomics/proteomics… and tissues and cells derived from transgenic rats and in particular transgenic mice, have also been overlooked
- some areas of reproductive/ developmental toxicity have not been covered…"

Also, the expert members of the former CSTEE (Scientific Committee on Toxicity, Ecotoxicity and the Environment) came to the conclusion that, for the foreseeable future, the use of live animals in toxicity testing remains essential in order to perform reliable risk assessments (Anon., 2004c).

Thus two expert committees of the Commission feel that, scientifically spoken, the development of alternatives has not yet reached the stage that one would tend to believe from the context of the current cosmetics legislation (2003/15/EEC, OJ L66, 11/03/2003) and coming chemicals legislation (Anon., 2003b). The approach of pushing legislation without having the necessary scientific knowledge appears to be quite threatening for the safety of cosmetics and their free circulation within and outside the EU market.

In order to support the replacement of *in vivo* by *in vitro* methods, research projects have been initiated within the Sixth Framework Research Programme of the EU, in particular for some of the fields where alternatives are still lacking at the moment (e.g. ReProTect, AcuteTox, Predictomics, Sens-it-iv and others). It must be emphasised, however, that these projects only started very recently and that the first results can only be expected around 2010.

Moreover, validation is a time-consuming activity that can only start after proper development of a relevant method and its prevalidation (Fentem and Balls, 1997). From the past, we know that it may take up to 8 years to have a method ready and implemented into the legislation. Finally, one must recognise that for certain endpoints, such as repeated dose toxicity testing, no acceptable proposals for *in vitro* development exist at this moment (Prieto et al., 2005).

One can argue that science advances quickly and proposals will emerge within the near future. Unfortunately, reality has shown differently. A literature search over the past 6 years (1998-2003) was carried out in Germany using standard databases including Chemical Abstracts, Medline, Embase, Biosis, SciSearch, Derwent Drug Pat., SCRIP Ring doc-ZEBET, ICCVAM (The Interagency Coordinating Committee on the Validation of Alternative Methods), and Monographs (Garthoff, 2003, 2005). It appeared that, of the apparently impressive number of articles published on the topic of 3R alternatives, most were reviews and only a limited number contained original research results (22 articles over 6 years).

Furthermore, most articles were not published in top scientific journals, which also handicapped their impact.

The case of repeated dose toxicity testing: Subacute and subchronic toxicity

Of relevance for cosmetic products and their ingredients are oral, dermal and inhalation tests, for which several OECD and EU-accepted *in vivo* tests are available today.

The currently indispensable oral repeated dose toxicity testing is still carried out in animals (rats and/or dogs) and, as mentioned before, to date no scientifically acceptable alternative methods exist (Eskes and Zuang, 2005). The current *in vivo* tests provide a so-called NOAEL (no observable adverse effect level) for the ingredient under consideration, a value that gives an estimation of the dosage that causes no adverse effects in the animal after repeated exposure (28 days or 90 days) (Anon., 2003b). This NOAEL is used to calculate the margin of safety (MoS) or uncertainty factor of the ingredient. The systemic exposure (SED) is taken into consideration by the following equation MoS= NOAEL/SED. According to the WHO, the calculated MoS for a given ingredient must be at least 100 in order to consider the ingredient to be safe for human use (Anon., 2003b). Thus the determination of a realistic NOAEL value has key importance in risk assessment.

When considering the replacement of the *in vivo* repeated dose toxicity test, it is enlightening to have a look at the endpoints of such a study in order to understand the complexity of its replacement. As reflected in tables 2 to 7, a large set of measurements and findings are involved. The total number of observations and measurements per test substance and per concentration can easily amount to more than a hundred.

If one aims to replace these complex *in vivo* tests by alternatives, one has to develop a testing strategy that provides the same level of knowledge as obtained *in vivo*. This implies an integrated approach, where *in silico* methodology needs to be combined with a battery of *in vitro* tests, and probably also refined *in vivo* tests, in order to include all relevant toxicological endpoints of different types of organs.

One can imagine that the physico-chemical and biological properties, the ADME (absorption, distribution, metabolism, excretion) and ADMET (ADME under toxicological conditions) data of molecules and classes of molecules (read-across strategy) can contribute (Coecke et al., 2005). Thus SAR and QSAR (quantitative structure activity relationship) and mechanism-based PK/PD (pharmaco-kinetic/pharmaco-dynamic models) analysis (Prieto et al., 2005) data certainly play a role in integrated testing, but their importance should not be overestimated either. They are mainly important as a first step for priority setting and further decision-making.

A whole range of *in vitro* tests such as cultures of primary cells and a variety of cell lines can provide important toxicological information on organs of different species (including man), in particular when these are combined not only with traditional

Tab. 2: Overview of the simultaneous estimation of different measurements and findings in subacute and subchronic toxicity testing, part "clinical observations"

clinical observations	- daily for overt signs and mortality
histopathology	- weekly for :
haematology	- skin and fur condition
body weight	- eyes
food consumption	- mucosa
organ weight	- respiratory function
clinical biochemical parameters	- circulatory function
urinanalysis	- nervous system function

Tab. 3: Overview of the simultaneous estimation of different measurements and findings in subacute and subchronic toxicity testing, part "histopathology"

clinical observations	- brain	- heart
histopathology	- spinal cord	- trachea & lungs
haematology	- pituitary	- aorta
body weight	- (para)thyroid	- gonads
food consumption	- thymus	- uterus
organ weight	- oesophagus	- accessory sex organs
clinical biochemical parameters	- salivary glands	- ♀ mammary gland
urinanalysis	- stomach	- prostate
	- small intestines	- urinary bladder
	- large intestines	- gal bladder
	- liver	- lymph nodes
	- pancreas	- peripheral nerve
	- kidneys	- bone marrow
	- adrenals	- skin
	- spleen	- eyes

Tab. 4: Overview of the simultaneous estimation of different measurements and findings in subacute and subchronic toxicity testing, part "haematology"

clinical observations	- haemoglobin concentration
histopathology	- erythrocyte count
haematology	- packed red blood cell volume (PCV)
body weight	- leukocyte count
food consumption	- platelet count
organ weight	- reticulocyte count
clinical biochemical parameters	- activated partial thromboplastin time
urinanalysis	

Tab. 5: Overview of the simultaneous estimation of different measurements and findings in subacute and subchronic toxicity testing, parts "body weight", "organ weight" and "food consumption"

clinical observations		
histopathology		
haematology		
body weight	- weekly observation	
food consumption	- 3d or 6d intervals (+ water consumption)	
organ weight	At necropsy	Relative weights
clinical biochemical parameters	- brain	- per body weight
urinanalysis	- liver	- per brain weight
	- kidneys	
	- lungs	
	- heart	
	- spleen	
	- thymus	
	- adrenal glands	
	- gonads	

Tab. 6: Overview of the simultaneous estimation of different measurements and findings in subacute and subchronic toxicity testing, part "clinical biochemical parameters"

clinical observations	- glucose
histopathology	- urea nitrogen
haematology	- creatinine
body weight	- total protein, albumin and globulin
food consumption	- bilirubin
organ weight	- phosphorus
clinical biochemical parameters	- Ca^{2+}, Na^{+}, K^{+}, Cl^{-}, HCO_3^{-}
urinanalysis	- aspartate aminotransferase
	- alanine aminotransferase
	- creatine kinase
	- sorbitol dehydrogenase
	- γ-glutamyl transferase
	- alkaline phosphatase
	- glutamate dehydrogenase
	- ornithine carbamyltransferase

Tab. 7: Overview of the simultaneous estimation of different measurements and findings in subacute and subchronic toxicity testing, part "urinanalysis"

clinical observations	- volume
histopathology	- appearance
body weight	- microscopy
haematology	- pH
food consumption	- osmolality
organ weight	- phosphorus
clinical biochemical parameters	- protein content
urinanalysis	- occult blood
	- glucose
	- ketones
	- bilirubin and urobilinogen
	- creatinine
	- α_{2u}-globulin
	- N-acetyl-β-D-glucoaminidase

technology but in particular with new "–omics" technology (Elaut et al., 2002). Indeed, a series of *in* vitro liver, kidney central nervous system, pulmonary, haematopoietic system preparations do exist (Prieto et al., 2005). Of these, liver cell cultures and, in particular, hepatocyte cultures, have been studied very extensively, but even these have not yet been validated or taken up into legislation (Coecke et al., 1999). Hepatocytes in culture, like other primary cells in general, dedifferentiate and loose their specific properties, making them less than ideal for long-term tests (Rogiers and Vercruysse, 1993; LeCluyse, 2001; Elaut et al., 2005).

Only recently an innovative technology was proposed to counteract dedifferentiation of primary cells in culture by inhibiting cell proliferation (Rogiers et al., 2004). The principle is based on the addition of histone deacetylase inhibitors to the perfusion solution during the isolation procedure of the cells and during their culture. These molecules are known as experimental anticancer drugs (Vanhaecke et al., 2004a). They interact with the acetylation/deacetylation process of histones in the cell nucleus (Lopez-Rodas et al., 1993; Magnaghi-Jaulin et al., 2000; Gregory et al., 2001). Hyperacetylated histones are formed, which result in DNA molecules becoming more accessible for transcription factors. Effects on the stabilisation of primary hepatocytes exposed to histone deacetylase inhibitors are available. Interesting results have been measured on cell cycle parameters (Papeleu et al., 2003, 2005, 2006), gap junctional intercellular communication (Vinken et al., 2006), apoptosis (Vanhaecke et al., 2004b), xenobiotic biotransformation (Henkens et al., 2005), and stem cell technology (Snykers et al., 2006). This new methodology is believed to be equally applicable to other primary cells (Rogiers et al., 2004) and research is ongoing.

Under the assumption that this innovative technology will result in the production of stable, functional primary cells of different organs, one still has to perform their full characterisation. Relevant markers and toxicological endpoints have to be developed for each of the toxicological targets. Furthermore, as it is known that *in vivo* all organs and tissues function in a coordinated and integrated way, it becomes crucial to also develop *in vitro* interrelationships between the different culture systems. Until now, the latter has been completely ignored.

What this example clearly shows is that innovation indeed takes place and opens perspectives, even in the context of a replacment strategy for repeated dose toxicity testing, but also that it follows its own rhythm and not that imposed by fixed timeframes.

Conclusions

The basic message is that scientific development continues as long as scientists are offered a positive environment for their work in terms of adequate funding, an up-to-date technological environment and a perspective for a decent carrier.

Innovation and break-through results, however, cannot be forced but are the normal outcome of continuous efforts and building up of knowledge by dedicated scientists. Thus scientific work follows its own rules and it seems absurd to try to impose legislative time deadlines as a way of putting pressure on the scientific community.

Efforts to speed up the outcome of results, however, can be obtained by a positive attitude towards opening up the field of alternatives by

Reorientation

(i) to attract top scientists to the field of alternatives,

(ii) to attract scientists of areas of science not yet involved in alternatives;

(iii) to stop focusing on cosmetics.

Rejuvenation

(iv) to provide more and better research opportunities for young researchers;

(v) to involve more young scientists in decision-making and strategy-building.

Requalification

(vi) to better coordinate and prioritise research projects on alternatives;

(vii) to stimulate scientists to invest in quality and to publish in scientific journals with a high impact factor.

Rather than political lobbying for non-realistic legislative deadlines and blocking the whole field, efforts could better be combined to stimulate a positive environment for good science.

References

Anonymous (2003a). COM (2003) 644 final. Proposal concerning the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH), (2003), http://europa.eu.int/eur-lex/en/com/pdf/2003/com2003_0644en.html (accessed 20/12/2004).

Anonymous (2003b). SCCNFP/0690/03 Final. Notes of Guidance for testing of Cosmetic Ingredients for their Safety Evaluation, adopted by the SCCNFP during the 25th plenary meeting of 20 October 2003.

Anonymous (2004a). Report for establishing the timetable for phasing out animal testing for the purpose of the cosmetic directive, issued by ECVAM, 30/4/2004.

Anonymous (2004b). SCCNFP/0834/04 Opinion concerning "Report for establishing the time-table for phasing out animal testing for the purpose of the cosmetics directive issued by ECVAM (30/4/2004), adopted by the SCCNFP on 1/7/2004 by means of the written procedure.

Anonymous (2004c). CSTEE 2004, Opinion of the Scientific Committee on toxicity, ecotoxicity and the environment (CSTEE) on the BUAV-ECEAE report on "the way forward-action to end animal toxicity testing". Adopted by the CSTEE during the 41st plenary meeting of 8/1/2004 [C7/VR/csteeop/anat/080104 D(04)].

Coecke, S., Blaauboer, B. J., Elaut, G. et al. (2005). Toxicokinetics and metabolism. In C. Eskes and V. Zuang (eds.), *ATLA 33 Suppl. 1*, 147-175.

Coecke, S., Rogiers, V., Bayliss, M. et al.. (1999). The use of long-term hepatocyte cultures for detecting induction of drug metabolism enzymes: the current status. *ATLA 27*, 597-638.

Elaut, G., Papeleu, P., Rogiers, V. and Vanhaecke, T. (2002). Practical aspects of in vitro biotransformation studies during early drug development. In S. G. Pandalai (ed.), *Recent*

Research and Development in Drug Metabolism and Disposition (167-198). Kerala, India: Transworld Research Network.

Elaut, G., Henkens, T., Papeleu, P. et al. (2005). Molecular mechanisms underlying the dedifferentiation of isolated hepatocytes and their cultures. *Curr. Drug Metabol.*, in press.

Eskes, C. and Zuang, V. (2005). Alternative (Non-Animal) methods for cosmetic testing; current status and future prospects. *ATLA 33 Suppl. 1*, 1-228.

Fentem, J. H. and Balls, M. (1997). The ECVAM approach to validation. In L. F. M. Van Zutphen and M. Balls (eds.), *Animal Alternatives. Welfare and Ethics-Developments in Animal and Veterinary Sciences* (1083-1089). Amsterdam, The Netherlands 27: Elsevier.

Garthoff, B. (2003). Alternative methods development-Pipeline: sufficient projects in face of EU Chemical White Paper and Cosmetic Legislation ? Report on 4th European Workshop of National Platforms on Alternative Methods, 28-30/11/2003, Brussels (http://ecopa.vub.ac.be, accessed 10/8/2005).

Garthoff, B. (2005). Alternatives to animal experimentation: The regulatory background. *Toxicol. Appl. Pharmacol. 207*, S388-S392.

Gregory, P. D., Wagner, K. and Hörz, W. (2001). Histone acetylation and chromatin remodelling. *Exp. Cell Res. 265*, 195-202.

Henkens, T., Vinken, M., Snykers, S. et al. (2005). Effect of Trichostatin A on maintenance of hepatic functions in primary cultures of rat hepatocytes. *Biochem. Pharmacol.*, submitted, 2005.

LeCluyse, E. L. (2001). Human hepatocyte culture systems for the in vitro evaluation of cytochrome P450 expression and regulation. *Eur. J. Pharm. Sci. 13*, 343-368.

Lopez-Rodas, G., Brosch, G., Georgieva, E. I. et al. (1993). Histone deacetylase. A key enzyme for the binding of regulatory proteins to chromatin. *FEBS Lett. 317*, 175-180.

Magnaghi-Jaulin, L., Ait-Si-Ali, S. and Harel-Bellan, A. (2000). Histone acetylation and the control of the cell cycle. *Prog. Cell Cycle Res. 4*, 41-47.

Papeleu, P., Loyer, P., Vanhaecke, T. et al. (2003). Trichostatin A induces differential cell cycle arrests but does not induce apoptosis in primary cultures of nitrogen-stimulated rat hepatocytes. *J. Hepatol. 39*, 374-382.

Papeleu, P., Coutant, A., Willaert, A. et al. (2006). Anti-proliferative and anti-apoptotic activity of the histone deacetylase inhibitor 4-Me2N-BAVAH in primary heptocytes. *J. Hepatol.*, submitted.

Papeleu, P., Vanhaecke, T., Elaut, G. et al. (2005). Differential effects of histone deacetylase inhibitors in tumor and normal cells – what is the toxicological relevance? *Crit. Rev. Toxicol. 35*, 363-378.

Pauwels, M. and Rogiers, V. (2004). Safety evaluation of cosmetics in the EU. Reality and challenges for the toxicologist. *Toxicol. Lett. 15*, 151(1):7-17.

Prieto, P., Clemedson, C., Meneguz, A. et al. (2005). Subacute and subchronic toxicity. In: Eskes C. and Zuang V. (eds), *ATLA 33 Suppl. 1*, 109-116.

Rogiers, V. and Vercruysse, A. (1993). Rat hepatocyte cultures and co-cultures in biotransformation studies of xenobiotics. *Toxicol. 82*, 193-208.

Rogiers, V., Snykers, S., Papeleu, P. et al. (2004). Differentiation of stem cells and stabilization of phenotypical properties of primary cells. International patent application number PCT/EP2004/012134, 1/11/2004-1/11/2005, Vrije Universtiteit Brussel, Department of Toxicology.

Snykers, S., Vanhaecke, T., Papeleu, P., et al. (2006). Robust MAPCs differentiation to hepatocyte-like cells by sequential exposure to cytokines reflecting development. *Development*, submitted.

Russell, W. M. S. and Burch, R. L. (1959). *The principles of Humane Experimental Technique*. Methuen and Co Ltd, London (reprinted by the Universities Federation for Animal Welfare UFAW, 1992, Potters Bar, Herts, UK).

Vanhaecke, T., Henkens, T., Kass, G.E.N. and Rogiers, V. (2004a). Effect of histone deacetylase inhibitor Trichostatin A on spontaneous apoptosis in various types of adult rat hepatocyte cultures. *Biochem. Pharmacol. 68*, 753-760.

Vanhaecke, T., Papeleu, P., Elaut, G. and Rogiers, V. (2004b). Trichostatin A-like hydroxamate histone deacetylase inhibitors as therapeutic agents: toxicological point of view. *Curr. Med. Chem. 11*, 1629-1643.

Vinken, M., Henkens, T., Vanhaecke et al. (2006). Trichostatin A enhances gap junctional intercellular communication in primary cultures of adult rat hepatocytes. *Toxicol. Sci.*, in press.

Correspondence to
Vera Rogiers
Dept. of Toxicology
Vrije Universiteit Brussel
Laarbeeklaan 103,
1090 Brussels, Belgium
e-mail: vrogiers@vub.ac.be

Workshop 5.15
In vitro metabolism: Applications in pharmacology and toxicology

Metabolic Activation for In Vitro Systems

Angelika Langsch and Heinz Nau
Department of Food Toxicology, University of Veterinary Medicine Hannover, Germany

Summary:
A significant portion of substances is not directly toxic, but must be activated to reactive metabolites. For most compounds it is unknown if metabolic activation plays a role. In vitro systems presently can only determine the toxicological potential of the parent drugs, and not that of potentially toxic metabolites. The development of metabolic systems which can be incorporated within a particular in vitro technique is therefore of high priority. Ideally, the in vitro system should express all relevant enzymatic activities, because it is a priori not known which enzyme(s) are involved. Also, the metabolic system must be compatible with the in vitro assay. Main emphasis is placed on liver preparations as the liver exhibits the highest amount and complexity of metabolic enzymes.

Keywords: metabolic activation, liver preparation, in vitro system

Introduction

In the planned REACH project about 30,000 available chemicals have to be registered, evaluated and authorised regarding their properties, including their toxic, carcinogenic, mutagenic, and teratogenic potential. The determination of all required data using *in vivo* experiments would mean the use of an immense number of animals. Hence, especially with regard to the concept of the 3Rs, *in vitro* methods are regarded to be a much wanted alternative.

However, a significant number of substances show toxic effects *in vivo* but not *in vitro*. Well known examples are polycyclic aromatic hydrocarbons (e.g. benzo(a)pyrene), cyclophosphamide, aflatoxin B1, Vitamin A and acrylamide. These substances are not directly toxic, but must be activated to reactive metabolites, which then exert their toxic potential by reacting with constituents of the cell, especially proteins and nucleic acids. Aflatoxin B1, for instance, is metabolised to aflatoxin-8,9-epoxide by CYP 3A4 and CYP 1A2 in a phase I reaction. This intermediate product can then be detoxified by a subsequent phase II reaction to the GSH-conjugate or undergo covalent binding to the DNA and thus be mutagenic (see fig. 1). Although for the given examples metabolic activation pathways are well established, for most compounds it is unknown whether metabolic activation plays a role.

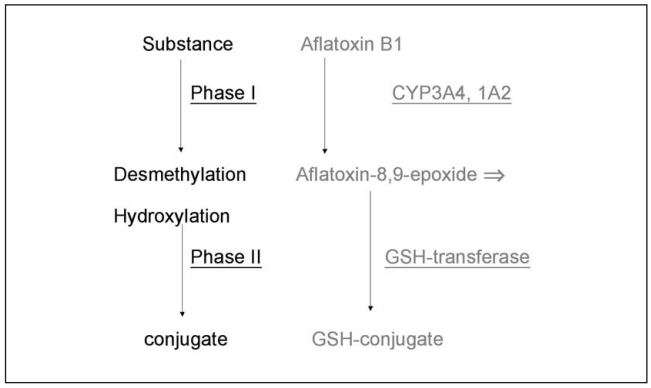


Fig. 1: Metabolic activation of aflatoxin B1.
Aflatoxin B1 is metabolised to aflatoxin-8,9-epoxide by CYP 3A4 and CYP 1A2 in a phase I reaction. This intermediate product can then be detoxified by the subsequent phase II reaction to the GSH-conjugate or undergo covalent binding to DNA, thus exerting its mutagenic effects.

To date *in vitro* systems can only determine the toxicological potential of the parent drugs and not that of potentially toxic metabolites. The development of metabolic systems which can be incorporated within a particular *in vitro* technique is therefore of high priority. Ideally, the metabolic *in vitro* system should express all the relevant enzymatic activities, because during screening of large numbers of substances it is *a priori* not known which enzyme(s) are involved. Also, the metabolic system must be compatible with the *in vitro* assay. The relevance of this approach consists in the adjustment of the existing *in vitro* systems to the *in vivo* situation. The aim is to reveal indirectly acting toxic or teratogenic substances which is presently not possible.

In this paper different target systems for teratogenic substances, as well as different metabolic activation systems, are introduced and their benefits and handicaps discussed. For exemplification the F9 test system is described in more detail.

Metabolic activation systems

Different metabolic systems appear suitable for such a task (Rueff et al., 1996), and the main emphasis is placed on liver preparations, as in most cases the liver exhibits the highest number and complexity of metabolic enzymes:

- S9 (liver 9.000 x g) preparation as used in the Ames test
- Hepatocytes or liver slices
- Genetically engineered cells expressing relevant metabolic enzymes

Subcellular liver fractions (S9-mix / microsomes)
These fractions are easy to handle "standard methods". S9-mix is used in the Ames test to detect genotoxic substances that need to be metabolically activated. However, for the present application many more mechanisms need to be considered than for the mutagenic response of bacteria. Hence, a number of possible limitations must be pointed out: all co-factors must be added. A substantial number of animals is necessary to obtain the subcellular liver fractions and differences have been analysed when subcellular and cellular systems were compared. Thus it is difficult to obtain physiologically relevant data. A further disadvantage when working with microsomes is given by the absence of most phase II enzymes. On the other hand, no additional (cell) culture is needed and the systems can be automated.

Hepatocytes and liver slices
Precision-cut liver slices are easy to prepare but exhibit only a short culture time. Hepatocytes can be obtained freshly isolated or cryopreserved. Both preparations cover all relevant pathways and contain all relevant enzymes and cofactors at physiological levels. Although liver slices and fresh cells can only be obtained directly after liver resection or removal, cryopreserved cells can be stored for long time periods and are thus always available. However, they show a decrease in phase II activity and phase I inducibility when compared to fresh cells. Another adverse property is the varying quality of hepatocytes and liver slices from different isolation experiments.

Genetically engineered cells
These cell lines, e.g. transfected V79, show two major advantages compared with hepatocytes: they are easy to handle and always available. However, they cover only a limited amount of enzymes, mostly phase I enzymes, which are not expressed at physiological levels.

Target systems

Embryonic Stem cell Test (EST)
For this assay, the pluripotent mouse stem cell line D3 is used. Under routine conditions the cells grow undifferentiated. If appropriate conditions are applied, they can e.g. differentiate into contracting cardiomyocytes. The cells are cultured with different concentrations of the test substance, and the number of contracting clones is monitored to determine an ID_{50} value (50% inhibition of differentiation). Also an IC_{50} (50% inhibition of

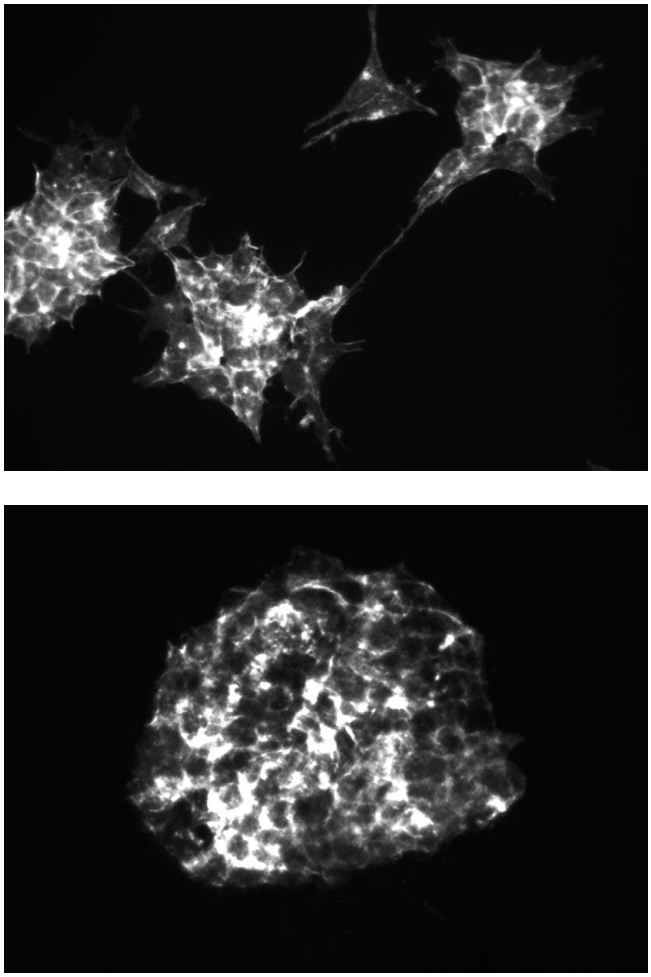


Fig. 2: Undifferentiated (bottom) and differentiated (top) F9 cells. Under routine conditions, F9 cells show an undifferentiated morphology and grow in clumps. However, differentiation to a neuronal-like morphology can be induced by teratogenic substances, e.g valproic acid (VPA). Blue: nucleus, green: cytoskeleton.

growth) is identified and compared with the IC_{50} of 3T3 Balb/c fibroblasts. Thus the teratogenic as well as the cytotoxic potential of the substance are determined (Spielmann et al., 1997).

Micro Mass culture (MM)
In this assay the limb bud or mid-brain cells of GD 12 mouse embryos are examined. They are cultured with the test substances and their development is monitored by Alcian blue stain. Teratogenic substances lead to morphologic changes in the target cells, exhibiting inhibition of differentiation and growth. As endpoints an ID_{50} (Alcian blue stain) and IC_{50} (cytotoxic effects resulting in reduction of neutral red uptake) are determined (Flint 1993).

Whole Embryo Culture (WEC)
For this assay whole rat embryos of GD 9.5 are cultured with the test substances in roller bottles. Different parameters are monitored as growth (crown-rump-length, yolk sac diameter, protein contents), differentiation (number of somites, morphology), and incidence of abnormalities. To classify no, weak and strong embryotoxic substances the IC_{NOEL} and IC_{max} of the embryos are compared to the IC_{50} of 3T3 fibroblasts (Piersma et al., 1996).

F9 cell line
The F9 cell line is derived from mouse testicular teratocarcinoma cells. Under routine conditions they show an undifferentiated morphology growing in clumps. However, differentiation to a neuronal-like morphology can be induced by teratogenic substances, e.g. valproic acid (VPA) (see fig. 2) (Göttlicher et al., 1998; Lampen et al. 1999). The changes taking place during differentiation closely mimic events of early mouse embryogenesis (Hogan et al., 1983).

VPA (2-n-propylpentatoic acid) is an antiepileptic drug particularly used for the treatment of several forms of epilepsy and bipolar disorders and as migraine prophylaxis (Sörensen,1988). Unfortunately, it proved to have considerable teratogenic potential in humans, inducing neural tube defects. A number of VPA analogues with varying teratogenic activity in mice have been synthesised which show that the teratogenic potential depends strictly on the structure of the compound (Nau et al., 1991). The teratogenic analogues bear a free carboxylic group and an α-hydrogen atom at the branching point of carbon atom C-2, but no further branching may be present (Nau and Löscher, 1986; Nau et al., 1991). Furthermore, the teratogenic effect of VPA and its derivatives is stereospecific. The (S)-enantiomer of 4-yn-VPA, for instance, is highly teratogenic in mice while the (R)-enantiomer shows little to no effect (Nau et al., 1991).

Remarkably, only *in vivo* teratogenic VPA-derivatives induce cell differentiation of F9 cells, whereas non-teratogenic VPA-derivatives do not show this effect (Lampen et al., 1999). The differentiation cannot only be monitored morphologically but also by the increased expression of differentiation markers. As F9 cells differentiate to a neuronal-like morphology, the expression of the NCAM-gene (neuronal cell adhesion molecule) is induced by teratogenic test substances (Lampen et al., 2005). This effect can further be demonstrated in a luciferase reporter gene assay coupled to viral promoter sequences. F9 cells are transiently transfected with a plasmid comprising the viral (RSV) promoter-driven luciferase gene. The promoter is only activated and thus the reporter gene only expressed by the teratogenic derivatives of VPA. Luciferase activity can therefore only be measured in F9 cells cultured with substances showing a teratogenic potency. The reporter activity highly corresponds to the teratogenic potential of the derivative (Lampen et al., 1999).

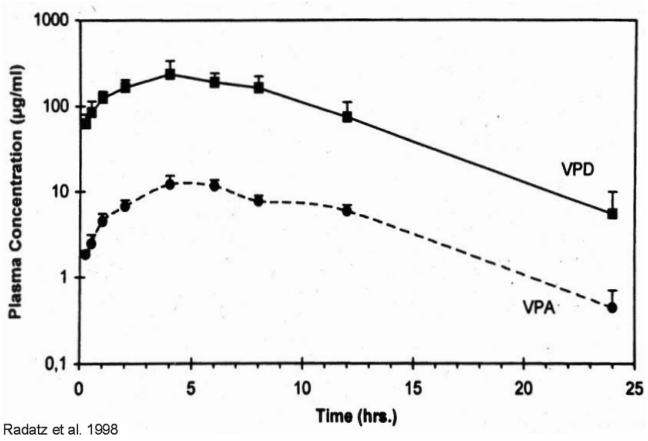


Fig. 3: Metabolic activation of valpromide (VPD) to valproic acid (VPA).
In vivo measurement of VPD metabolism and VPA formation in NMRI mice treated with 3 mmol/kg BW VPD (Radatz et al., 1998).

Combinations

In contrast, two other analogues, valpromide (VPD) and 2-pentyl-4-pentynoic hydroxamic acid, which exhibit a teratogenic potency *in vivo*, cannot be detected with the F9 cell system. Both of these substances need to be metabolically activated: VPD to VPA (see fig. 3) (Radatz et al., 1998) and the hydroxamic acid to its corresponding carbonic acid.

First experiments have been done using S9-mix as metabolic activation system. Two approaches can be distinguished: direct addition of S9-mix to the F9 cells and pre-incubation of S9-mix with the test substances followed by incubation of the F9 cells with the extracted medium (see fig. 4). Direct addition showed two problems: a high amount of S9 is toxic to the F9 cells, but with less S9 too little metabolite is formed to induce differentiation of the target cells. Hence, direct addition is not suitable for the given system. For the pre-incubation, in contrast, no limit of the applicable amount of S9-mix is observed. VPD metabolism with human, rat (Aroclor induced) and mouse S9-fractions was compared. Rat S9 metabolised the highest proportion of VPD but, as *in vivo*, VPA is further metabolised and not available for differentiation of the F9 target cells. Human S9 and mouse S9 metabolise similar proportions of VPD, but with human S9 more VPA is found than with mouse S9 (see tab. 1). The formed VPA is first extracted to be concentrated and subsequently added to F9 culture medium to induce differentiation. Due to the cyto-

toxic potential of the S9-mix further studies to improve this procedure need to be done. However, these first *in vitro* results incorporating the metabolic activation correspond well with the *in vivo* results. For the given experiment, human S9 is best suited to mimic the *in vivo* situation and eliminate interspecies differences.

Further studies to test the well known proteratogen cyclophosphamide were done combining metabolic activation systems with target systems, e.g. hepatocytes with chicken brain cultures (Bruinink et al., 2002), or whole embryo culture (Ozolins et al., 1995), or genetically engineered cells (V79) with the D3 stem cell line (Bremer et al., 2002). Bruinink et al. (2002) used primary hepatocytes to metabolise cyclophosphamide and isophenphos and tested the neurotoxic effect of the metabolites with embryonic chicken brain cells. The cells were able to discriminate between the toxicity of the parent drugs and their metabolites and between metabolites with an unspecific cytotoxic activity (cyclophosphamide) and metabolites with a high potential to damage specific nerve cell populations (isophenphos). Ozolins et al. (1995) co-cultured murine whole embryos with primary maternal hepatocytes and exposed them to cyclophosphamide. They showed that the hepatocytes were necessary for the expression of embryotoxicity, which was concentration-dependent. Bremer et al. (2002) cultivated the D3 stem cell line with supernatants of the genetically engineered mammalian cell line V79, transfected with CYP 2B1 cDNA to metabolise cyclophosphamide. The combined system was able to detect the embryotoxic potential in a reporter gene assay for developmental cardiotoxicity.

Tab. 1: *In vitro* measurement of VPD metabolism and VPA formation after pre-incubation with S9-mix.
Rat S9 metabolised the highest proportion of VPD but, as *in vivo*, VPA is further metabolised and not available for differentiation of the F9 target cells. Human S9 and mouse S9 metabolise similar proportions of VPD, but with human S9 more VPA is found than with mouse S9.

S9-mix	VPD metabolism	VPA formation
human	30-45%	25-35%
rat (Aroclor induced)	50-80%	10-20%
mouse	30-50%	10-15%

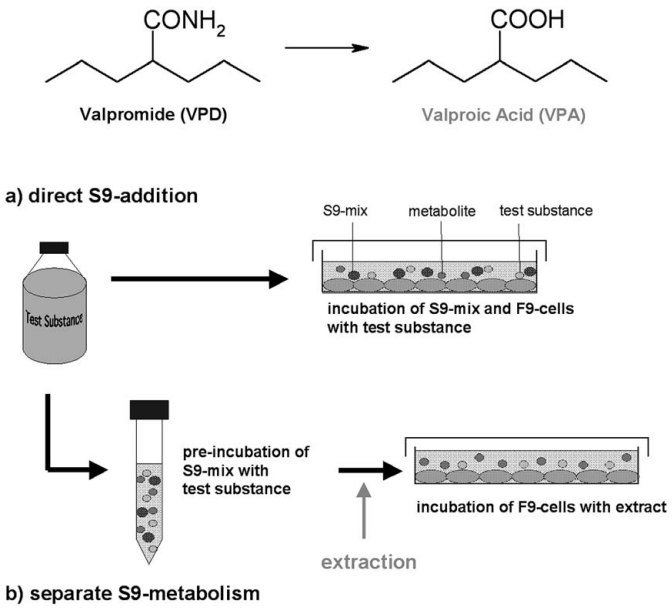


Fig. 4: Metabolic activation using S9-mix.
Two approaches can be distinguished: direct addition of S9-mix to the F9 cells and pre-incubation of S9-mix with the test substances followed by incubation of the F9 cells with the extracted medium.

Discussion and conclusion

A reliable *in vitro* system for the identification of the embryotoxic potential of substances should reproduce all aspects of embryogenesis. Therefore, also the maternal metabolism and foeto-placental interaction have to be taken into account, because some proteratogens require bioactivation to provide active molecules that interact with the developing embryo. The described examples show that the current approaches of combining existing *in vitro* assays with metabolic activation systems are feasible to extend the usefulness of the *in vitro* test procedures. The addition is necessary to exclude false negative results regarding adverse effects when characterising a substance. However, substantial differences in the results were observed in the same indicator system when cellular or subcellular hepatic activating systems were used or when subcellular preparations from rats and humans were compared (Rueff et al., 1996). Ideally, the metabolic system should express all relevant enzymes, because during screening of many substances it is *a priori* not known which enzyme(s) may be involved. Also, the needed cofactors should be present in physiological concentrations. Generally, intact cells of liver origin seem to satisfactorily mimic the metabolic activation that occurs *in vivo* and they reflect the *in vivo* genotoxicity better than S9-mix (Rueff et al., 1996).

The possible advantages of using hepatocytes for activation are several: (1) human cells can be used to avoid species differences; (2) they contain enzymes and cofactors at physiological levels; (3) high activity of both phase I and phase II enzymes is present. These metabolic activation systems are now being developed for incorporation into *in vitro* systems for development of robust testing systems which can be transferred to other laboratories. To be practicable they have to be simple to handle, reproducible and cost-effective.

The relevance of these approaches consists in the adjustment of the existing *in vitro* systems to be more reflective of the complex *in vivo* situation with the aim to reveal indirectly acting toxic or teratogenic substances which is presently not possible.

References:

Bremer, S., Pellizzer, C., Coecke, S. et al. (2002). Detection of the embryotoxic potential of cyclophosphamide by using a combined system of metabolic competent cells and embryonic stem cells. *Altern. Lab. Anim. 30*, 77-85.

Bruinink, A., Yu, D. and Maier, P. (2002). Short-term assay for the identification of neurotoxic compounds and their liver derived stable metabolites. *Toxicol. In Vitro 16*, 717-724.

Flint, O. P. (1993). In vitro tests for teratogens: desirable endpoints, test batteries and current status of the micromass teratogen test. *Reproductive Toxicol. 76*, 383-395.

Göttlicher, M., Werling, U., Siehler, S. et al. (1998). Cellular actions of valproic acid and its teratogenic derivatives: Activation of peroxisome proliferators activated receptors (PPARs) and differentiation of teratocarcinoma cells. *Naunyn Schmiedebergs Arch.Pharmacol. 358*, R775.

Hogan, B. L. M., Barlow, D. P. and Tilly, R. (1983). F9 teratocarcinoma cells as a model for the differentiation of parietal and visceral endoderm in the mouse embryo. *Cancer Surv. 2*, 115-140.

Lampen, A., Siehler, S., Ellerbeck et al. (1999). New molecular bioassays fort he estimation of the teratogenic potency of valproic acid derivatives in vitro: activation of the peroxisomal proliferator-activated receptor (PPARdelta). *Toxicol. Appl. Pharmacol. 160*, 238-249.

Lampen, A., Grimaldi, P. A. and Nau, H. (2005). Modulation of peroxisome prolierator-activated receptor delta activity affect NCAM and ST8Sia IV (PST1) induction by teratogenic VPA-analogues in F9 cell differentiation. *Mol. Pharmacol. 68*, 193-203.

Nau, H. and Löscher, W. (1986). Pharmacologic evaluation of various metabolites and analogs of valproic acid: Teratogenic potencies in mice. *Fundam. Appl. Toxicol. 6*, 669-676.

Nau, H., Hauck, R. S. and Ehlers, K. (1991). Valproic acid-induced neural tube defects in mouse and human: Aspects of chirality, alternative drug development, pharmacokinetics and possible mechanisms. *Pharmacol. Toxicol. 69*, 310-321.

Ozolins, T. R. S., Oglesby, L. A., Wiley, M .J. and Wells, P. G. (1995). In vitro murine embryotoxicity of cyclophosphamide in embryos co-cultured with maternal hepatocytes: development and application of a murine embryo-hepatocyte co-culture model. *Toxicology 102*, 259-74.

Piersma, A., Bechter, R., Krafft, N. et al. (1996). An interlaboratory evaluation of five pairs of teratogens on postimplantation rat embryo culture. *ATLA 24*, 201-209.

Radatz, M., Ehlers, K., Yagen et al. (1998). Valnoctamide, valpromide and valnoctic acid are much less teratogenic in mice than valproic acid. *Epilepsy Res. 30*, 41-48.

Rueff, J., Chiapella, C., Chipman, et al. (1996). Development and validation of alternative metabolic systems for mutagenicity testing in short-term assays. *Mutation Res. 353*, 151-176.

Sörensen, K. V. (1988). Valproate: A new drug in migraine prophylaxis. *Acta Neurol. Scand. 78*, 346-348.

Spielmann, H., Pohl, I., Doehring, B. et al. (1997). The embryonic stem cell test: an in vitro embryotoxicity test using two permanent mouse cell lines – 3T3 fibroblasts and embryonic stem cells. *In Vitro Toxicol. 10*, 119-127.

Acknowledgements

Supported by grants of the EU (ReProTect), the BMBF and EU-RTN-HPRN-CT-2002-00370. We thank Andrea Seiler and Horst Spielmann for cooperation with D3 stem cells as well as the members of the "Metabolic Activation Group" of ReProTect.

Correspondence to

Prof. Dr. Dr. h.c. Heinz Nau
Stiftung Tierärztliche Hochschule Hannover
Institut für Lebensmitteltoxikologie und Chemische Analytik
Lebensmitteltoxikologie – Gebäude 115 Abt. 44
Bischofsholer Damm 15
30173 Hannover, Germany
e-mail: Heinz.Nau@tiho-hannover.de

Workshop 5.16
Reproductive toxicology – the EU ReProTect project

Placental Perfusion – A Human Alternative

Tina Mose and Lisbeth Ehlert Knudsen
Institute of Public Health, University of Copenhagen, Denmark

Summary
Foetal exposures to environmental and medicinal products have impact on the growth of the foetus (e.g. cigarette smoke) and development of organs (e.g. methylmercury and Thalidomide). Perfusion studies of the human term placenta enable investigation of placental transport of chemical substances between the mother and foetus. Dual perfusion of a single cotyledon in the human placenta can contribute to a better understanding of the placental barrier, transport rate and mechanisms of different substances and placental metabolism. The perfusion system has recently been established in Copenhagen and represents a supplement and alternative to animal testing, bypassing the animal to human extrapolation.
Placentas are readily obtainable from most births upon informed consent from the mothers and are considered a promising tissue alternative/supplement to animal experiments. The system is validated as a part of work package 2 of the integrated project ReProTect.

Keywords: human placenta, foetal exposure, human alternative, biomonitoring

Introduction

Foetal growth and development is extremely dependent on placenta functionality, as the major role of the placenta is to exchange oxygen, nutrients, hormones and waste products between the mother and the foetus. In addition, the placenta has an important role in hormone synthesis and metabolism of compounds that regulate and maintain pregnancy (Ahokas and Anderson, 1987). The placenta develops from embryonic tissue and by the end of the third gestation week embryonic blood circulates through the capillaries of the villi (Khong and Pearce, 1987). The barrier function of the placenta membrane depends on molecular size, polarity, lipid solubility and protein binding of the substances. Both active transport mechanisms and facilitated transport have been described (Syme et al., 2004). Substances able to cross the placenta have the potential to cause adverse effects directly on foetal development or placenta function. Substances not able to cross the placenta may harm the foetus through effects on the mother, e.g. toxicity.

The first evidence of reproductive toxicity caused by a foetal exposure from maternal intake was the thalidomide disaster in 1957-1961. Pregnant women from approximately 46 countries worldwide were prescribed thalidomide as a safe sedative. Given between the 34th and 50th day of pregnancy, thalidomide may exert teratogenic effects seen as skeletal malformation especially of the limbs. These malformations were reported in more than 10,000 babies who survived the exposure. When investigated further, the mechanism of action was specific to humans, but was later confirmed in a second animal species. As a result of this disaster the demands on testing of drugs to be used during pregnancy were increased (Botting, 2002; Brent, 2004) .

In the early 1970s it was reported that prenatal alcohol exposure can cause mental retardation, facial malformations, prenatal and/or postnatal growth retardation (Riley and McGee, 2005; West and Blake, 2005). In 1971 the effect of the synthetic non-steroidal estrogen diethylstilbestrol (DES) on the reproductive system became evident. The drug was prescribed to prevent mis-

carriage and other pregnancy complications but, as an unknown teratogenic effect, it also caused carcinomas in the vagina and cervix in young women and malformation of reproductive organs in girl and boy offspring (Palmlund, 1996; Swan, 2000). This led to a broader definition of the concept of reproductive effects, including not only functional and cognitive effects seen at birth, but also effects seen later in life caused by foetal exposure. It is now well established that maternal smoking (Brown et al., 1988; Habek et al., 2002; Philipp et al., 1984) and maternal exposure to methyl mercury (Hamada et al., 1997), lead (Banks et al., 1997) and environmental chemicals such as polychlorinated biphenyls (Schantz, 1996; Tilson et al., 1990) may cause adverse developmental defects in the offspring.

Tools for foetal risk assessments

Risk characterisation of foetal development from maternal exposure to chemicals and pharmaceuticals is most commonly based on results from animal studies, sometimes supported by *in vitro* studies or more seldom by epidemiological studies. The OECD test guidelines for reproductive toxicity testing include a prenatal developmental study (TG 414), one-generation study (TG 415), two-generation study (TG 416) and a reproduction/developmental toxicity screening test (TG 421 + TG422) (http://ecb.jrc.it/testing-methods/, 2005). These studies are designed to provide dose-response relationships concerning the toxic effect of prenatal exposure on the pregnant test animal, on the developing organism in the uterus, and the effect on the integrity and performance of the male and female reproductive systems after one or two exposed generations, see table 1. Fertility, growth, malformations, and survival are typical effect endpoints in the offspring after administration of very high doses to the pregnant dam. The studies are designed to induce toxicity but not mortality in the pregnant animals at the highest dose level. The observed foetal toxicity is either caused by a direct developmental effect or a maternally mediated effect. Analysis of the distribution of compounds in and between the maternal and foetal compartments is not required in all instances. However, applying the OECD guideline regarding toxicokinetics (B.36), where pregnant animals are sometimes requested, may provide such information (tab. 1).

In animal testing, the foetal exposure is not investigated in detail, as no confirmatory studies on the presence or distribution of compound in foetal tissue are required. An adverse effect in the foetus is considered an indication of foetal exposure. Human placental perfusion can provide information about transplacental transfer, placental metabolism, storage, acute toxicity and potential role of transporters, vascularisation and foetal exposure. It is beneficial to use human placenta tissue, as extrapolations from animal to human are bypassed. This is very important as the human placenta is unique in structure and only resembles placenta from certain monkey species like baboons (Enders and Blankenship, 1999). The human placenta is hemomonochorial where the foetal tissue is in direct contact with maternal blood and villous where the placenta is divided into small vascular

Tab. 1: OECD guidelines on reproductive toxicity testing

Relevant OECD guidelines	N	Aim	Dosing	Endpoints
Prenatal developmental study (TG 414)	20	Malformations	Only during organogenesis	Live and dead foetuses, resorptions, foetal weight, skeletal anomalies, delayed ossification, organ anomalies
One-generation study (TG 415)	20	Fertility and pre-post natal development	Before and during mating period and for females during gestation and lactation	Fertility of males and females, birth and litter size, growth and survival of offspring, histopathology. Pre-and perinatal death and malformations – only as smaller litters
Two-generation study (TG 416)	20	Fertility and pre-post natal development	Before and during mating period and for females during gestation and lactation. Continued dosing of offspring	Fertility of males and females, birth and litter size, growth and survival of offspring, histopathology, sexual maturation, oestrus cyclicity, semen quality. Pre-and perinatal death and malformations – only as smaller litters. Cover all periods.
Reproduction/ developmental toxicity screening test (TG 421, 422)	8	Screening	Two weeks before mating to postnatal day 4	Fertility, birth and litter size, growth and survival until PND 4, and histopathology in paternal animals. TG 422: haematology. Pre-and perinatal death and malformations – only as smaller litters
Toxicokinetics (B36)	4	Absorption, distribution, excretion and metabolism of substance.	Single or repeated dose, animals sacrificed at different times after exposure.	Amount of substance in urine, excreta, bile, plasma, and milk. Distribution: whole body autoradiographic techniques or quantitative analysis of substance and/or metabolites in tissues and organs.

N = number of pregnant animals per dosing and control group.
TG 422 = combined repeated dose toxicity study with the reproductive/developmental toxicity screening test.

units, i.e. cotyledons (Faber et al., 1992; Leiser and Kaufmann, 1994). The use of human tissue in placenta perfusion can help overcome the differences in kinetics, placenta structure, sensitivity, duration of gestation and background levels. It carries potential to replace or reduce the number of animals used for toxicological testing. Results from the human placenta perfusion test system may provide important knowledge on transplacental transfer of new chemical substances as well as on environmental exposures of hazardous compounds in humans.

The aim of this presentation is to describe the placenta perfusion method established at the University of Copenhagen to study the transplacental transfer and placental storage of environmental compounds. The system allows us to collect, and determine the concentration of compound in umbilical cord blood, maternal and paternal blood samples, providing useful samples for biomonitoring studies in families.

It is also possible to determine endpoints such as placental metabolism, placental transporters, and the presences of specific biomarkers in blood and placenta samples. Placental perfusion will be developed in the integrated projects ReProTect and NewGeneris. The data on placental transfer will be included in another ReProTect workpackage exploiting data for QSAR consideration (Hareng et al., 2005).

Placenta perfusion

The placenta perfusion system was first described and developed by Panigel and later modified by Schneider and other research groups to enable more systematic studies (Panigel, 1985; Schneider and Huch, 1985). Several research groups have subsequently used similar system to investigate the placental transfer and metabolism of different compounds, especially pharmaceutical drugs as sulindac, oxcarbazepine, lamotrigine and theophylline administered to the maternal compartment (Ala-Kokko et al., 1997; Lampela et al., 1999; Myllynen et al., 2001; Myllynen et al., 2003; Nanovskaya et al., 2002; Omarini et al., 1992; Pienimaki et al., 1995).

Materials and methods

Chemicals

Antipyrine 98% was purchased from Aldrich-Chemie (Steinheim, Germany), p-acetopenetidide 97% (phenacetin) from Acros organics (Geel, Belgium), dextran from Leuconostoc mesenteroides from Sigma-Aldrich (Steinheim, Germany), and heparin 5000 IE/ml from SAD (Copenhagen, Denmark). Potassium chloride, potassium dihydrogen phosphate, calcium chloride ($CaCl_2 \cdot H_2O$), magnesiumsulphate ($MgSO_4 \cdot 7H_2O$) and sodium hydrogen carbonate were manufactured by Merck (Damstadt, Germany). Sodium chloride was from J. K. Baker (Deventer, Holland) and hydrochloric acid 6N from Bie & Berntsen (Roedovre, Denmark). D(+)-Glucose monohydrate, phosphoric acid 85%, and methanol were purchased from Applichem (Biochemia, Darmstadt, Germany). All solvents were of gradient grade used for high-performance liquid chromatography.

Instruments

The maternal and foetal circulation is driven by two essential roller pumps (Watson Marlow SciQ 323E/D). Another roller pump (MasterFlex easy-load 3 model 77800-50) serves to retain buffer in the perfusion chamber. The perfusion chamber is handmade of Plexiglas material at the University of Copenhagen. Tygon R3603 tubing is used (Saint-Gobain). The lidded reservoirs are placed on common stirring devices. The tissue is homogenised by a Turrax homogeniser (T25 basic, 8mm). An eight canal Powerlab System (ADInstruments, Oxfordshire, UK) is connected to two flow-thru oxygen electrodes (Microelectrodes, New Hampshire, USA), a pH meter, and a temperature amp from ADInstruments and an on-line computer software programme (see fig. 1).

Logistics

The placentas are obtained from births delivered at the Danish National Hospital by elective Caesarean section. The weekday before delivery, the Elective Section Team at the hospital, composed of midwife, obstetrician and anaesthesia nurse, meets with the pregnant woman. On occasion of this meeting, information about the placenta perfusion study is given and written material is handed out. The material contains a short summary of the study, a detailed declaration of informed consent, a questionnaire on maternal data and exposure, and a folder from the ethics committee about patient rights when participating in scientific experiments. Further information on the study is available on the homepage (www.pubhealth.ku.dk/placenta/). If the pregnant woman wants to participate, a signed declaration of informed consent is given and a questionnaire is filled out and returned before hospitalisation for Caesarean section. Coordi-

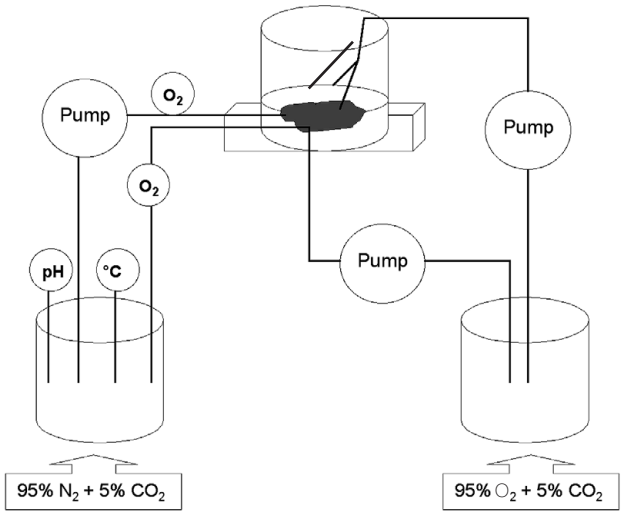


Fig. 1: Schematic presentation of the Placenta perfusion system

nation with the midwife makes it possible to collect placentas immediately after birth.

Placentae

Immediately after delivery the placenta is weighed. Umbilical blood is drawn into a VenoJect sodium heparin tube. 25 ml freshly prepared 37ºC warm *Krebs-Ringer buffer (NaCl, 121 mM; $NaHCO_3$, 24.9 mM; KCl, 3.0 mM; KH_2PO_4, 0.4 mM; $CaCl_2$, 1.3 mM; $MgSO_4$, 0.58 mM) with added heparin (25,000 IU/l) and glucose (9 mmol/l), adjusted to pH 7.2-7.4 with HCl, is injected into umbilical arteries and vein. The placenta is carried to the laboratory in an insulated box. Upon arrival, the placenta is inspected and remarks on size, ruptures, infarcts, and condition in general are noted.

Placental perfusion method

The dual recirculation placenta perfusion system at the University of Copenhagen was, with modifications, adapted from the University of Oulu, Finland (Myllynen et al., 2003). Briefly, a foetal artery-vein pair supplying a well-defined cotyledon is cannulated using Luer stub adapters of size 20G and 18G, respectively. *Krebs-Ringer buffer is manually injected into the foetal circulation, the cotyledon is isolated and connected to the pump system to ensure that inflow equals outflow. A tissue sample from beside the removed cotyledon is taken, cut into small pieces, homogenised and stored at -20ºC. The cotyledon is placed in the perfusion chamber with the maternal side up and two Luer stub adapters (size 18G) are gently inserted into the intervillous space 5-7 mm below the maternal surface. The perfusion chamber enables the cotyledon to be wetted in maternal perfusate with a minimum of evaporation due to the small entrances for tubing representing arteries end veins. The cotyledon is preperfused for 30 minutes to restore adequate oxygen to the tissue and to ensure stable foetal venous outflow. Antipyrine and compound is added on the maternal side and samples are drawn from both reservoirs after 0, 15, 30, 60, 90, 120, and 150 min. The volume collected is replaced with perfusion buffer.

The perfusion buffer in both the maternal and foetal reservoir is *Krebs-Ringer buffer (200ml) with 8.5 g/l and 30 g/l dextran added, respectively. The foetal perfusate is constantly gassed with 95% N_2/ 5% CO_2 and the maternal perfusate is gassed with 95% O_2/ 5% CO_2. The perfusion flow rate is 3.5 ml/min in the foetal circulation and 12 ml/min in the maternal circulation.

After the end of perfusion the cotyledon is weighed, cut into small pieces, homogenised, and stored at -20ºC. The volume of perfusion liquid in the maternal and foetal chambers is measured. Samples are placed on ice and within 30 minutes the blood cells are precipitated by centrifugation at 4,000 x g for 5 minutes. The supernatant is stored at -20ºC until analysis; 0.2 ml is needed for the antipyrine analysis.

Safety concerns

No information is *a priori* available about the health status of the pregnant woman and all tissues are considered potentially infected with e.g. hepatitis and HIV. Thus, laboratory personnel handling placentas are vaccinated with Hepatitis B and all procedures are performed with personal protection including gloves, glasses and laboratory coats.

Analysis of antipyrine

200 µl supernatant is added to 200 µl H_3PO_4 (0.5 M) with phenacetin added (10 µg/ml) as internal standard. Antipyrine and phenacetin are analysed using a reverse phase LaChrom HPLC system equipped with an L-7100 pump, an L7200 autosampler, an D-7000 interface, a L-7300 column oven and a L-7400 UV detector (Merck, Hitachi). The stationary phase is a C18 column (NUCLEOSIL C-18, ODS, 20 x 4,6 mm, 5 µm particles) with a SecurityGuard precolumn (Phenomexes C-18, ODS, 4 mm L x 3.0 mm ID). The mobile phase is a degassed methanol/water (45/55 v/v) solution adjusted to a flow rate of 1 ml/min. Injection volume is 25 µl, oven temperature is 30-32ºC and absorbance is detected at 254 nm. A calibration graph of antipyrine (0, 0.5, 1.5, 10, 25 µg/ml) and phenacetin is constructed with *Krebs Ringer as matrix.

Ethics

Placenta perfusion studies cause minimal ethical problems, partly because the experiments are non-invasive – causing no harm to mother or child - and partly because placentas are normally discarded and incinerated after birth. The use of placenta and umbilical cord blood in scientific research requires a signed and informed consent from the mother. To obtain maternal and paternal blood, a similar declaration is required from both the mother and the father. Much time must be scheduled to inform participants properly and straightforwardly about their contribution to the study and the study in general.

Results

Antipyrine

Antipyrine crosses the placenta by passive diffusion; therefore it is a good functionality marker of the placental perfusion system. In figure 2, the transfer of antipyrine from the maternal to the foetal compartment is shown. Within 150 minutes, 40 µg/ml antipyrine diffused from maternal to foetal chamber and an equilibrium between maternal and foetal circulation was almost established. In figure 3, the feto-maternal ratio (FM) of the transplacental passage from the same experiment is shown. The FM ratio is the concentration ratio of antipyrine in foetal and maternal perfusate.

Controlling physiological experimental conditions

When perfusions comply the following demands, they are evaluated as successful:

-Foetal venous outflow stable within 7.0 ± 0.2 ml per 2 min.
-Volume loss from foetal circulation < 20 ml after end perfusion.
-The FM ratio of antipyrine transfer > 0.7 after 150 minute of perfusion.
-The transfer of oxygen from maternal to foetal circulation has to be sufficient pO2 (foetal vein) >> pO2 (foetal artery).
-Temperature in foetal perfusate 37 ± 2°C.

-pH in foetal perfusate within 7.2-7.4 ± 0.1.
-Time from birth of child to first cannulation < 30 minutes.

Conclusions

The human placenta perfusion system is a reliable and feasible supplement to the existing animal tests used for human foetal risk assessments, as this test system gives useful information on placental transfer, placental storage and metabolism in the placenta tissue, without the need for extrapolation to a different species. Results from human perfusion studies can improve the human foetal exposure assessment, providing important information on placental toxicokinetics. Such data are seldom included in reproductive toxicity studies in animals. The human placental perfusion system is a supplement to animal reproductive testing to be further developed within the integrated project ReProTect.

References

Ahokas, R. A. and Anderson G. D. (1987). The Placenta as an Organ of Nutrition. In j.Patrick Lavery (ed.), *The Human placenta Clinical Perspectives* (207-220). New York: Aspen Publishers.

Ala-Kokko, T. I., Pienimaki, P., Lampela, E. et al. (1997). Transfer of clonidine and dexmedetomidine across the isolated perfused human placenta. *Acta Anaesthesiol. Scand. 41(2)*, 313-319.

Banks, E. C., Ferretti, L. E. and Shucard, D. W. (1997). Effects of low level lead exposure on cognitive function in children: a review of behavioral, neuropsychological and biological evidence. *Neurotoxicology 18(1)*, 237-281.

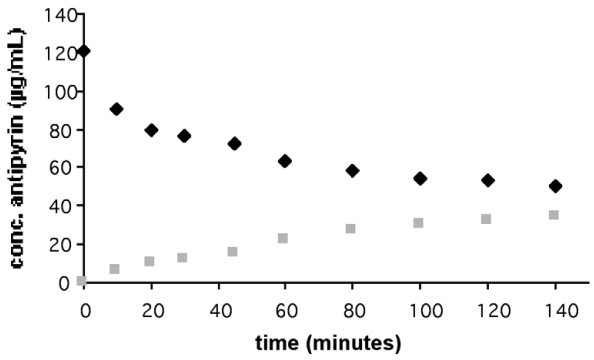


Fig. 2: Transplacental transfer of antipyrine

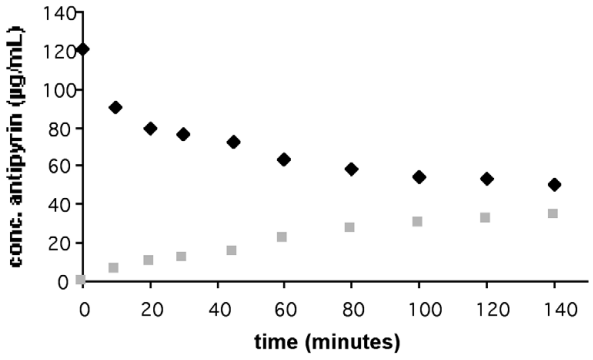


Fig. 3: FM ratio of antipyrine transfer

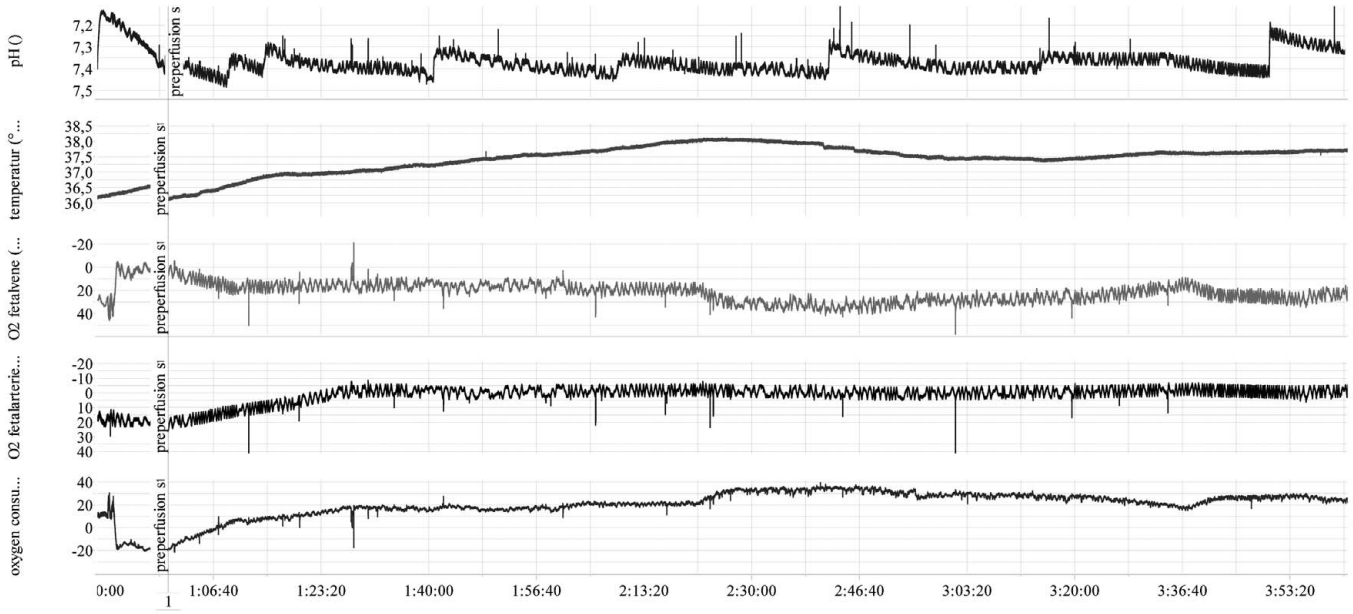


Fig. 4: Powerlab data from a placenta perfusion

Botting, J. (2002). The History of Thalidomide. Drug News *Perspect. 15(9)*, 604-611.

Brent, R. L. (2004). Utilization of animal studies to determine the effects and human risks of environmental toxicants (drugs, chemicals, and physical agents). *Pediatrics 113 (4 Suppl)*, 984-995.

Brown, H. L., Miller, J. M., Jr., Khawli, O. and Gabert, H. A. (1988). Premature placental calcification in maternal cigarette smokers. *Obstet. Gynecol. 71(6 Pt 1)*, 914-917.

Enders, A. C. and Blankenship, T. N. (1999). Comparative placental structure. *Adv. Drug Deliv. Rev. 38(1)*, 3-15.

Faber, J. J., Thornburg, K. L. and Binder, N. D. (1992). Physiology of Placental-Transfer in Mammals. *American Zoologist 32(2)*, 343-354.

Habek, D., Habek, J. C., Ivanisevic, M. and Djelmis, J. (2002). Fetal tobacco syndrome and perinatal outcome. *Fetal Diagn. Ther. 17(6)*, 367-371.

Hamada, R., Arimura, K. and Osame, M. (1997). Maternal-fetal mercury transport and fetal methylmercury poisoning. *Met. Ions. Biol. Syst. 34*, 405-420.

Hareng, L., Pellizzer, C., Bremer, S. et al. (2005). The integrated project ReProTect: a novel approach in reproductive toxicity hazard assessment. *Reprod. Toxicol. 20(3)*, 441-452.

Khong, T. Y. and Pearce J. M. (1987). Development and Investigation of the Placenta and Its Blood Supply. In J. Patrick Lavery (ed.), *The Human placenta Clinical Perspectives* (25-33). New York: Aspen Publishers.

Lampela, E. S., Nuutinen, L. H., Ala-Kokko, T. I. et al. (1999). Placental transfer of sulindac, sulindac sulfide, and indomethacin in a human placental perfusion model. *Am. J. Obstet Gynecol. 180(1 Pt 1)*, 174-180.

Leiser, R. and Kaufmann, P. (1994). Placental Structure – in A Comparative Aspect. *Experimental and Clinical Endocrinology 102(3)*, 122-134.

Myllynen, P. K., Pienimaki, P. K. and Vahakangas, K. H. (2003). Transplacental passage of lamotrigine in a human placental perfusion system in vitro and in maternal and cord blood in vivo. *Eur. J. Clin. Pharmacol. 58(10)*, 677-682.

Myllynen, P., Pienimaki, P., Jouppila, P. and Vahakangas, K. (2001). Transplacental passage of oxcarbazepine and its metabolites in vivo. *Epilepsia 42(11)*, 1482-1485.

Nanovskaya, T., Deshmukh, S., Brooks, M. and Ahmed, M. S. (2002). Transplacental transfer and metabolism of buprenorphine. *J. Pharmacol. Exp. Ther. 300(1)*, 26-33.

OECD test guidelines on reproductive toxicity testing (2005). http://ecb.jrc.it/testing-methods/.

Omarini, D., Barzago, M. M., Aramayona, J., Bortolotti, A., Lucchini, G., and Bonati, M. (1992). Theophylline transfer across human placental cotyledon during in vitro dual perfusion. *J. Med. 23(2)*, 101-116.

Palmlund, I. (1996). Exposure to a xenoestrogen before birth: the diethylstilbestrol experience. *J. Psychosom. Obstet. Gynaecol. 17(2)*, 71-84.

Panigel, M. (1985). Past, present, and future of placental perfusion experiments. *Contrib. Gynecol. Obstet. 13*, 132-136.

Philipp, K., Pateisky, N., and Endler, M. (1984). Effects of smoking on uteroplacental blood flow. *Gynecol Obstet Invest. 17(4)*, 179-182.

Pienimaki, P., Hartikainen, A. L., Arvela, P. et al. (1995). Carbamazepine and its metabolites in human perfused placenta and in maternal and cord blood. *Epilepsia 36(3)*, 241-248.

Riley, E. P. and McGee, C. L. (2005). Fetal alcohol spectrum disorders: an overview with emphasis on changes in brain and behavior. *Exp. Biol. Med. (Maywood.) 230(6)*, 357-365.

Schantz, S. L. (1996). Developmental neurotoxicity of PCBs in humans: what do we know and where do we go from here? *Neurotoxicol. Teratol. 18(3)*, 217-227.

Schneider, H. and Huch, A. (1985). Dual in vitro perfusion of an isolated lobe of human placenta: method and instrumentation. *Contrib. Gynecol. Obstet. 13*, 40-47.

Swan, S. H. (2000). Intrauterine exposure to diethylstilbestrol: long-term effects in humans. *APMIS 108(12)*, 793-804.

Syme, M. R., Paxton, J. W. and Keelan, J. A. (2004). Drug transfer and metabolism by the human placenta. *Clin. Pharmacokinet. 43(8)*, 487-514.

Tilson, H. A., Jacobson, J. L. and Rogan, W. J. (1990). Polychlorinated biphenyls and the developing nervous system: cross-species comparisons. *Neurotoxicol. Teratol. 12(3)*, 239-248.

West, J. R. and Blake, C. A. (2005). Fetal alcohol syndrome: an assessment of the field. *Exp. Biol. Med. (Maywood.) 230(6)*, 354-356.

Acknowledgements

The studies are supported by a national grant from the Danish Research Centre for Environmental Health (0-302-02-9/5), CHILDRENGENONETWORK (QLK4-CT-2002-02198), and a PhD stipend for Tina Dam Mikkelsen from the Danish National Graduate School of Public Health (GRASPH) and ReProTect (LSHB-CT-2004-503257). The authors gratefully acknowledge the collaboration with the Maternity Unit and Dr. Morten Hedegaard at Rigshospitalet, Copenhagen.

Correspondence to

Ass. Prof. Ph.D. Lisbeth E. Knudsen
Center for Health and Society, 5.2.12
Øster Farimagsgade 5
1014 Copenhagen K
Denmark
e-mail: l.knudsen@pubhealth.ku.dk

Theme 6
Modelling

Session 6.1
QSAR approach: Acceptance and implementation

The Use of *In Silico* Technologies to Predict Toxicity and Fate: Implementation and Acceptance

Mark T. D. Cronin
School of Pharmacy and Chemistry, Liverpool John Moores University, Liverpool, England

Summary
In silico technologies provide the capability to predict the toxicity and fate of chemicals without the use of animals. These approaches include the use of existing data, read-across, (quantitative) structure-activity relationships (Q)SARs and expert systems. There are a large number of implementation activities relating to the use of (Q)SARs in the forthcoming REACH legislation. Much work is being focussed on developing strategies for testing, user-friendly tools and determining how a prediction from a (Q)SAR may be evaluated for acceptance. This paper reviews these areas and identifies key areas for future research effort.

Keywords: QSARs, expert systems, REACH, validation

Introduction and aims

In silico technologies encompass a large number of approaches that may assist in the prediction of toxicity and fate of new and existing chemicals. The aims of this paper are to introduce briefly the concept of *in silico* techniques for toxicity prediction, evaluate how and where they may be used for regulatory purposes, their implementation and acceptance criteria, and finally point out key areas for further research. This paper has been written with particular emphasis on regulatory usage of these approaches. It should also be noted that many of the techniques described below have considerable potential use in product development. Finally, whilst every effort will be made to keep this article up-to-date, in such a fast moving area, there will inevitably be some redundancy of the information even with expeditious publication.

In silico technologies

There has been an appreciation that biological activity is related to chemical structure since the mid-nineteenth century, and, from the end of that century, that toxic potency may be related to physico-chemical properties (Schultz et al., 2003). The last century has seen a maturing of this science into what are now frequently termed "*in silico*" methods for the prediction of toxicity and fate. As a term of reference for this paper, *in silico* is taken to imply a broad group of techniques including the use of existing data, read-across (also known as analogue or category approaches), quantitative structure-activity relationships ((Q)SARs) and expert systems. The reader should be aware that others may not use such expansive terminology and may take *in silico* to cover fewer of these topics. *In silico* approaches are of interest as they allow predictions of toxicity and fate from chemical structure alone, thus supporting the reduction in animal testing. For a full overview of the development and use of *in silico* methods the reader is referred to Cronin and Livingstone (2004).

Should (acceptable) data for a particular chemical for an endpoint of interest already exist, there should be no need to predict or measure toxicity. There are many sources of toxicity and fate data for use, the most accessible of which have been reviewed recently by Cronin (2005). There are many places in the literature and on the internet containing toxicological information, one of the best starting places being the TOXNET online

database [http://toxnet.nlm.nih.gov/] (Cronin, 2005). There are also several commercial databases of toxicological information of varying cost and quality. Regulatory data could provide a very amenable source of information for risk assessment, although many of these databases are not generally openly available. Associated to the regulatory data are those being generated for the assessment of the High Production Volume (HPV) chemicals. Other sources of data are, of course, in-house repositories, although again these will not be freely available for the foreseeable future. As well as providing a resource for regulatory assessment, existing data may also be suitable to assist in model development and evaluation.

Whilst there are many existing data available, there are also a number of concerns about their use. The most fundamental concerns address the quality and relevance of the data, and whether they are even relevant for the decision being made, i.e. are they acquired according to an acceptable or reliable protocol, etc. Other issues relate to the practicalities of retrieving existing data, especially from historical databases, that may be paper-based, or stored on incompatible electronic databases. Efforts to record toxicity data onto freely available and widely compatible formats with mark-up languages (such as xml) should be encouraged. Furthermore there is the ever-present problem of the release, or use of, commercially sensitive data. As noted above, the current dogma is that many such data will not be released voluntarily, and due to the cost and effort involved in organising such data, it may not be possible to do so practically. There is also the complex problem of who will host, maintain and update databases of toxicological information. Despite these problems, there are two illustrations of what may be done. The Distributed Structure-Searchable Toxicity (DSSTox) Database Network has been developed as a project of the US Environmental Protection Agency's (EPA) Computational Toxicology Program [http:// www.epa.gov/nheerl/dsstox/]. The VITIC toxicological database is being developed by Lhasa Ltd. [http://www.lhasalimited.org/].

Read-across, category and analogue approaches are based on the assumption that "similar" compounds will have "similar" properties. Thus, if a compound is "similar" to a known toxic compound, the probability of it being toxic will be high. In some ways this can be considered to be a simplification of the traditional use of structure-activity relationships (SARs), although the mechanistic nature of the SAR need not be defined formally in read-across. There is effort to develop at least one tool (the US EPA's Analog Identification Method) to search databases and perform read-across automatically. It should be noted that this tool is not publicly available at the time of preparation of the manuscript, but is undergoing beta-testing.

(Q)SARs attempt to relate the toxicity of a series of chemicals to their physico-chemical and structural properties. In theory, models of some form can be developed for most toxic endpoints (i.e. qualitative or quantitative) using any of a wide variety of molecular descriptors and statistical techniques (Cronin, 2004a). (Q)SARs offer the capability to make a prediction from chemical structure alone and have the advantage that they may shed some light on the mechanism of toxic action. Models may be standalone, e.g. some form of regression equation, or be automated into a formal expert system. Expert systems may be defined as being "any formalised system, not necessarily computer-based, which enables a user to obtain rational predictions about the toxicity of chemicals" (Dearden et al., 1997). There are many approaches and philosophies to develop expert systems, and they have been applied to predict toxicity and fate by regulatory agencies worldwide (see the next section). There are a large number of expert systems available, both as freely available software and commercial systems; these are well reviewed by a number of authors including Combes and Rodford (2004).

Regulatory applications of *in silico* technologies

There is considerable use of (Q)SARs and other *in silico* methods to predict toxicity and fate by Regulatory Agencies worldwide. The area in which (Q)SAR has been used by regulatory agencies is very well covered and reviewed (Cronin et al., 2003a, 2003b, 2004b). It is generally accepted that *in silico* methods will be used for:

- prioritisation of existing chemicals for further testing or assessment
- classification and labelling of new chemicals
- risk assessment of new and existing chemicals

Very briefly, *in silico* technologies have been used by Health Canada in relation to the Domestic Substances List; by the Danish Environmental Protection Agency to develop a (Q)SAR database and support system; by the Japanese National Institute of Technology and Evaluation (NITE) to examine targets under the current Chemical Substances Control Law (CSCL); extensively in the United States by the Environmental Protection Agency (e.g. EPISuite, OncoLogic, ChemSTEER, E-FAST, PBT Profiler, Analog Identification Method) and by the Food and Drug Administration's (FDA) Center for Drug Evaluation and Research (CDER).

In the European Union the current focus of much effort in *in silico* toxicology is towards the ultimate implementation of the Cosmetics Directive and the Registration, Evaluation and Authorisation of CHemicals (REACH) legislation. Whilst there is much international effort to use (Q)SARs to predict toxicity and fate, the emphasis of the remainder of this manuscript will be on the use of *in silico* methods for the forthcoming REACH legislation.

Implementation of *in silico* technologies

There appear to be clear benefits of the use of *in silico* tools to make predictions within the framework of the REACH legislation. In particular, *in silico* methods play a clear role in Integrated Testing Strategies for toxicological endpoints (Combes et al., 2003; Worth and Balls, 2002). Existing data, as well as predictions from read-across and (Q)SARs, are well established as part of the strategy. Whilst much work has been done and is, of course, on-going, integrated strategies are still required for some toxicological endpoints. In addition, more consideration will be required as to the validation of these strate-

gies, as well as the individual components within them.

The implementation of *in silico* methods (particularly in REACH) will be linked strongly to issues of the acceptance of the methods and the evaluation of methods and any predictions from them. At the time of preparation of this paper, there are a number of international implementation activities on-going and planned. These include the various REACH Implementation Projects (RIPs) and RIP 3.3 in particular. There are also a number of industry-sponsored initiatives (e.g. CEFIC-LRI projects) as well as (non-industry) projects being organised through the European Chemicals Bureau at Ispra, Italy. A number of projects are being funded by European Union Framework and Non-Framework programmes as well as a number of national governmental initiatives in member states. No specific details of these projects are provided as many are on-going or are due to commence shortly. The interested reader should contact the relevant funding bodies for more details.

Acceptance

The acceptance of predictions from *in silico* toxicological methods is a complex matter. There is a clear role for the characterisation and evaluation and possible validation of *in silico* models. With regard to this issue, Annex IX of the REACH Proposal states:

"results obtained from valid (Q)SARs may indicate the presence or absence of a certain dangerous property. Results of (Q)SARs may be used instead of testing when the following conditions are met:

- results are derived from a (Q)SAR model whose scientific validity has been established
- results are adequate for the purpose of classification and labelling and risk assessment
- adequate and reliable documentation of the method is provided"

Further details are available from: http://europa.eu.int/comm/enterprise/reach/overview.htm.

The quest for methods to assess "scientific validity" has resulted in a number of activities to assist in the development of a scheme to "validate" (Q)SARs. At this point it should be noted that the (Q)SAR community has historically used the term "validation" in a somewhat different sense, to indicate the statistical quality and robustness of a model. This has, at times, caused confusion in the (Q)SAR community with the broader meaning of "validation" in the regulatory sense.

A set of criteria for the validation of (Q)SARs has been agreed upon by the Organisation for Economic Co-operation and Development (OECD), the so-called "OECD Principles". These state that a (Q)SAR should be associated with the following information:

- a defined endpoint
- an unambiguous algorithm
- a defined applicability domain
- appropriate measures of goodness-of-fit, robustness and predictivity
- a mechanistic interpretation, if possible.

The results of a pilot study to evaluate these criteria are available at: http://www.oecd.org/document/23/0,2340,en_2649_34365_33957015_1_1_1_1,00.html and, at the time of preparation of the manuscript, four "validation" studies are on-going, co-ordinated by the European Chemicals Bureau.

When one begins to attempt to validate a (Q)SAR, it soon becomes apparent that a considerable number of challenges must be faced. For instance, there is sparse availability of toxicity data for test sets to allow for a true assessment of predictivity. There are also considerations for commercial models that may not wish to have their models made transparent. It is also important that any validation process is not seen as a competition. There should be no winners in this activity, merely an appreciation of the merits (or otherwise) of individual predictions.

The validation process has therefore become somewhat controversial, with some calls for complete validation (Balls and Combes, 2005). To broaden the debate, it should be remembered that in terms of Annex IX quoted above, "valid" means valid for a specific purpose. It can be construed that the amount of information necessary to demonstrate validity should be context-dependent. This can be considered in relation to the acceptability of making a wrong decision and the acceptability of not being able to make a decision at all, due to lack of experimental data. Validation exercises (such as those being undertaken at the time of preparation of this paper) may be better served if they focus on characterising (Q)SARs according to the OECD principles as opposed to considering a model as being "validated" or "non-validated". In other words, if no experimental data are available, a prediction may have some value, providing the limitations of the model and the prediction are appreciated and understood. Thus, for instance, it may be appropriate to consider some level of confidence in a prediction of activity within the framework of the OECD principles.

Guidance Document(s) are required in order to facilitate the use and validation of (Q)SARs. OECD Guidance Documents on (Q)SAR Validation and Case Studies on regulatory acceptance of (Q)SARs are being planned and executed. Other activities include the development of a (Q)SAR Application Toolbox to assist the user in the application of *in silico* techniques. These and other requirements are discussed below.

Future directions and needs

Whilst a number of *in silico* tools are already available for use and are being used for regulatory purposes, there is still a requirement for on-going development of models, tools and guidance. Some of this development is planned and ongoing and much is funded at an international level (see, for instance the implementation section). With so much effort being placed in the development of *in silico* tools, there is a clear need for co-ordination of these efforts. Much is currently being performed and further leadership and planning will be required and is underway. There are specific issues to be considered, for instance, the availability of toxicity and fate data for modelling and model evaluation. There is recognition of the need for more data, but few definitive ideas of where these data will be sourced,

how to assess their consistency and reliability, how to perform quality assessment and assurance, storage and handling, and how to deal with commercially sensitive data. Other issues with regard to the application of *in silico* tools relate to the degree of validation that may be required for a model for any particular purpose; the formal definition and consideration of the applicability domain of a model; and how different predictive approaches may be considered in terms of consensus modelling. Education of model users (i.e. from regulatory agencies) and appropriate guidance documents (as described above) is also an area that will need much effort, and probably much good will from the trainers. There are, of course, many other equally pressing needs for the development of good *in silico* tools for the prediction of toxicity and fate. Few, if any, are insurmountable, and it is positive to see that much is currently being planned, performed and achieved at the international level to resolve all these issues.

Conclusions

In silico technologies cover a wide variety of approaches to predict toxicity and fate. Many of these are relevant to regulatory agencies. The REACH legislation has focussed the thoughts and efforts of the scientific community to address and resolve the issues faced by this "incomplete" area of science. Implementation of these approaches, in terms of providing user-friendly tools, education and acceptance criteria, is still required at many levels. Ultimately, modellers (developers and users) must accept that the use of predicted values is not an easy, quick and simple option, and pragmatism will be required.

References

Balls, M. and Combes, R. (2005). The need for a formal invalidation process for animal and non-animal tests. *ATLA 33*, 299-308.

Combes, R. D., Barratt, M. D. and Balls, M. (2003). An overall strategy for the testing of chemical for human hazard and risk assessment under the EU REACH systems. *ATLA 31*, 7-19.

Combes, R. D. and Rodford, R. A. (2004). The use of expert systems for toxicity prediction: illustrated with reference to the DEREK program. In M. T. D. Cronin and D. J. Livingstone (eds), *Predicting Chemical Toxicity and Fate* (193-204). Boca Raton FL, USA: CRC Press.

Cronin, M. T. D. (2004a). Predicting chemical toxicity and fate in humans and the environment – an introduction. In M. T. D. Cronin and D. J. Livingstone (eds), *Predicting Chemical Toxicity and Fate* (3-13). Boca Raton FL, USA: CRC Press.

Cronin, M. T. D. (2004b). The use by governmental regulatory agencies of quantitative structure-activity relationships and expert systems to predict toxicity. In M. T. D. Cronin and D. J. Livingstone (eds), *Predicting Chemical Toxicity and Fate* (413-13). Boca Raton FL, USA: CRC Press.

Cronin, M. T. D. (2005). Toxicological information for use in predictive modelling: quality, sources, and databases. In C. Helma (ed.), *Predictive Toxicology* (93-133). Boca Raton FL, USA: Taylor and Francis.

Cronin, M. T. D., Jaworska, J. S., Walker, J. D., et al. (2003a). Use of QSARs in international decision-making frameworks to predict health effects of chemical substances. *Environ. Health Persp. 111*, 1391-1401.

Cronin, M. T. D. and Livingstone, D. J. (2004). *Predicting Chemical Toxicity and Fate*. Boca Raton FL, USA: CRC Press.

Cronin, M. T. D., Walker, J. D., Jaworska, J. S., et al. (2003b). Use of QSARs in international decision-making frameworks to predict ecologic effects and environmental fate of chemical substances. *Environ. Health Persp. 111*, 1376-1390.

Dearden, J. C., Barratt, M. D., Benigni, R. et al. (1997). The development and validation of expert systems for predicting toxicity. *ATLA 25*, 223-252.

Schultz, T. W., Cronin, M. T. D., Walker, J. D. and Aptula, A. O. (2003). Quantitative structure-activity relationships (QSARs) in toxicology: a historical perspective. *J. Mol. Struct. (Theochem) 622*, 1-22.

Worth, A. P. and Balls, M. (2002). Alternative (non-animal) methods for chemicals testing: current status and future prospects. A report prepared by ECVAM and the ECVAM Working Group on Chemicals. *ATLA 30 Suppl. 1*, 1-125.

Acknowledgments

Far from being my own ideas, the thoughts presented in this paper are the culmination of hours of discussion over many years with a multitude of excellent scientists across the world. I'm certain we don't agree all the time, but I do thank them all for the intellectual stimulation they have provided.

Correspondence to

Mark Cronin
School of Pharmacy and Chemistry
Liverpool John Moores University
Byrom Street
Liverpool L3 3AF
England
e-mail: m.t.cronin@ljmu.ac.uk

Roles for QSAR in Risk Assessment

Gilman D. Veith
International QSAR Foundation to Reduce Animal Testing, Two Harbors, Minnesota, USA

Summary
Early pioneers in quantitative structure-activity relationship (QSAR) development believed in the premise that laboratory experiments should not be performed without a firm expectation of the results before going into the laboratory. The value of QSARs in risk assessment will be to provide us with those expectations for a wide variety of exposure and hazard assessment endpoints before the decision to require specific testing is made. Most of the current QSAR models are limited only by the lack of designed databases; however, as they evolve, QSAR models for most endpoints will undoubtedly be used to provide us with test expectations for thousands of untested chemicals. In so doing, QSAR will complement the 3Rs with a powerful new tool to minimise animal testing that is not likely to influence regulatory decisions. If it is true that 95% industrial chemicals have a lower probability to be classified as an endocrine disrupting chemical (EDC) than n-butyl aniline, avoiding testing on those chemicals can be achieved by QSAR screening. With the recent development of computer simulators of metabolic activation (e.g. the virtual liver), improved QSAR models for skin sensitisation, respiratory irritation and genotoxicity will follow quickly. Finally, the integration of QSAR models with in vitro methods holds great promise for the prudent use and interpretation of our testing and assessment resources.

Keywords: QSAR, in silico, REACH, non-animal, model domain

Introduction

Studies of quantitative structure-activity relationships (QSAR) can be traced back in the literature for more than a century. Whenever QSAR has been applied to the more complex problems in medicinal chemistry and toxicology, the models have had to be greatly simplified in order to fit the computational capabilities of the times. The publication of simplified models has often led to misapplication of the model as well as fuelling criticism of QSAR.

This millennium has ushered in vast new computational capabilities for QSAR. Just as has happened with other areas of science incorporating modern computing, it is tempting to describe QSAR in new terms such as "*in silico*", which make it appear to be a new specialised field or technology. It is important for us to remember, however, that QSAR is not a discipline or a field of research. Rather, it remains a scientific approach that can be applied to many scientific fields to understand the determinants of forecasting results of experimental methods involving chemicals before the experiments are performed. Together with international concurrence on the principles for validation of QSAR models, the next decade promises to bring unparalleled application of QSAR in risk assessment.

The evolution of the Registration, Evaluation and Assessment of CHemicals (REACH) legislation in Europe may offer *in silico* methods a central role in fulfilling data requirements over the next decade. The principles of validation of QSAR were articulated in the Setubal Principles (Jaworska et al., 2003), which have, after slight change in emphasis, become OECD principles for the validation of QSAR. One of the themes in the principles of validation for QSAR models is transparency. It is generally understood that, for QSAR models to be accepted by the regulatory community, transparency in terms of availability of the data and of the algorithm for the model is sought. This level of transparency is largely one of documentation, sufficient to allow the model to be reproduced by others, a basic tenet of science in general. Perhaps less frequently emphasised is the transparency of the mechanisms that underpin the model. Transparency in this mechanistic context is the extent to which the model can explain why the estimated endpoint value was estimated as it was and how a change in structure might affect the results.

Before describing some of the important roles for QSAR in regulatory risk assessment, it is important to review the basic elements of the QSAR approach. The first element of QSAR is to select a well-defined endpoint of chemical or biological activity to be modelled and to compile a database for chemicals that have been tested. The second element of QSAR is to compile a list of molecular descriptors for the structure of the chemical in the database. The third element of QSAR is to apply statistical methods to explore the variances in the data and identify useful relationships between the molecular descriptors and the endpoint of activity. The forth element of QSAR is to formulate a hypothesis about the observed relationship and modify the database in a way that will test the hypothesis. Many transparent QSAR models include the first three elements of the QSAR approach, but the majority of QSAR models offer scant records of efforts to redesign the database to reveal underlying mechanisms.

The lack of transparency regarding structural requirements for applying a QSAR leads to what I will call the "domain conundrum". Without a mechanistic basis for a model, it is difficult to

select chemicals that can really validate the model or explain the reasons the model can be used for a specific chemical in a regulatory application. To escape this conundrum in validation, the QSAR model domain is often defined as a multidimensional parameter space derived from the molecular descriptors in the training set. Defining a domain in these terms is an implicit argument that all chemicals that fall within that structure space should be expected to conform to the model, which is seldom correct and can easily be shown to be invalid.

QSAR in a nutshell

The essence of QSAR is the fundamental belief in chemistry that expects similar chemicals to behave similarly. As such, the essential role for QSAR is to serve as a guide in the selection of chemicals for testing of mechanistic understanding and not the after-the-fact modelling of data on a small number of chemicals. Like any model, QSAR can improve our hypotheses about why some chemicals are active and some are not active, or at least make the formulation of hypotheses more efficient. Formulating hypotheses and then selecting new chemicals for testing is the scientific method, and the only way to create a robust model with clear boundary conditions.

The basic steps in building a QSAR model for an iterative process begin with the compilation of a database for the endpoint for which the model is intended. Molecular descriptors are calculated for each chemical represented in the database and exploratory statistical methods are used to identify potential relationships. From the analysis of outliers, hypotheses are generated to examine those relationships by adding new chemicals to or subdividing the database. The revised training set is again analysed statistically and the process repeated. The hypotheses generated in the refinement of the model form the basis for defining the model domain in mechanistic terms. If one takes a dataset of biological effects, calculates 300 parameters for each of the chemicals and develops a statistical relationship, the resulting "model" represents only a part of one QSAR development step. The domains of such models must be described only as boundaries in a parameter space for the molecular descriptors that were statistically important in the model.

Roles for QSAR in regulatory applications

The goal of QSAR is not to produce a series of models to be used in place of laboratory tests, but rather to improve both the design and strategic use of test methods. QSAR is a perspective with which one can approach complex problems with techniques to distinguish important problems. QSAR improves hypothesis generation in research involving chemicals and, in so doing, tends to improve the experimental design in many area of scientific specialisation.

The short-term role for QSAR in regulatory risk assessment will be to add clarity to existing processes and make them more efficient. In OECD countries, the emphasis for QSAR will be to assist in the development of chemical categories based on common chemical behaviour. When categories are formed as analogues, the opportunity exists to extend the test data for any members of the categories to all members of the categories through methods of read-across. In developing categories both in classification and risk assessments, information is needed regarding the metabolites of the chemicals. Empirical data for metabolism is sparse and piecemeal, and QSAR-based metabolic simulators can provide information to strengthen these processes. Also in the short-term, there are sufficient QSAR models mature enough to use them to begin establishing testing and assessment priorities.

In the longer-term, which will allow time to improve QSAR models for important endpoints, the highest priority will likely be the use of QSAR to estimate missing values for screening information data set (SIDS) endpoints. In addition, a number of studies are demonstrating that QSAR models can be integrated with bioinformatics and high throughput testing to improve the efficiency of these capabilities. Combining all these uses of QSAR, the most significant role for QSAR will be to create a hypothesis-driven sequential testing strategy, which will eliminate all of the battery testing in risk assessment that is not ultimately used in the risk assessment (Bradbury et al., 2004). Whether these approaches are called intelligent testing strategies or 3rd generation risk assessment paradigms for chemicals, QSAR models will be central to making the logical decisions for testing requirements.

Historical barriers to QSAR

In addition to the point made earlier that predicting many forms of chemical activity was a larger computational problem than could be realistically approached with computers in the past, there are three barriers to QSAR modelling which have limited its development. The first of these has been chemical speciation and conformational analysis. Most QSAR models use molecular descriptors computed for parent chemicals even though different species are known to exist in natural systems. As stereoelectronic descriptors come into use, the gas phase minimum energy conformation of the chemical is commonly used, even though many conformations for flexible molecules would be expected at ambient conditions. As the sensitivity of QSAR models to conformational assumptions becomes more apparent, the importance in understanding the influence of conformation will grow. Fortunately, methods are now available to remove the historical conformational barrier.

The second historical barrier to QSAR modelling arose from the fact that the parent chemical structure is often not the actual toxicant causing the effect. Like many *in vitro* tests methods, QSAR models could not account for metabolic activation of chemicals. Mixing chemicals that were directly active with those that must be changed to toxic metabolites in a common training set usually creates outliers in modelling. Within the past five years, the first generation of metabolic simulators, which can convert parent chemicals to a plausible family of metabolites, have been developed. These computerised simulators of metabolic processes are limited only by the amount of high qual-

ity, systematic data on metabolism and are likely to be refined greatly over the next five years.

Since running a QSAR model both on a parent chemical and its 30-50 plausible metabolites can usually be accomplished "on the fly", what was a historical barrier will soon be a routine check for possible metabolic activation for specific models.

The third historical barrier to QSAR modelling has been the modelling of chemical reactivity and reactive toxicity. Chemical reactivity is important to model because it represents a series of mechanisms by which chemicals bind to cellular targets and cause toxic effects at low concentrations. Chemical reactivity is difficult to predict because the reactivity is due, in part, to inducible properties of chemicals as well as the more static properties of chemicals. The stereoelectronic structure of an isolated chemical can change significantly when it is in water or near membranes, proteins and DNA. Developing a formalism for modelling the inducible properties of chemicals to estimate what molecular interactions are most probable remains a challenge for QSAR modellers. Because reactive toxicity includes so many important biological effects for risk assessment and spans both genotoxic and non-genotoxic effects, it is a major gap in the ability of QSAR to respond to the need of risk assessment.

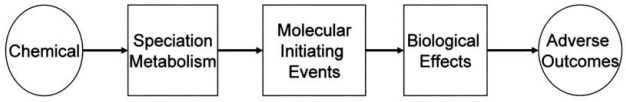
Framework for reactive toxicity models

Reactions of electrophiles with these nucleophiles give rise to one important distinction between reactive toxicity and non-reactive toxicity. Reactive toxicity encompasses the biological effects and measurable adverse outcomes resulting from an irreversible reaction of a xenobiotic chemical with endogenous molecules including proteins, nucleic acids, and lipids. While reactive toxicity includes protein and/or lipid oxidation, conjugation reactions, and the broad spectrum of covalent reactions, we have narrowed the breadth of possible reactions to begin illustrating use of the reactive toxicity framework to two-electron covalent reactions.

Molecular interactions between foreign chemicals and cellular components can initiate or disrupt cellular/organism processes and lead to a wide array of adverse outcomes, including acute failure of energy flow and nerve function, skin irritation/sensitisation, immune system dysfunction, reproductive and developmental impairment, idiosyncratic organ failure and death, mutagenicity and carcinogenicity. Clearly, since chemicals from different classes can cause similar biological effects, a single chemical can cause multiple biological effects, and many chemicals are "activated" by metabolic processes so the parent chemical is only a "carrier" and not the ultimate toxicant, a conceptual framework that clarifies the many controlling factors is needed.

All of these controlling factors, whether they be biochemical, cellular, organ or systems-based, converge at the point where a molecular initiating event occurs. Consequently, the reactants as well as the reactions involved in those molecular initiating events form the centrepiece of the conceptual framework presented in this work. In figure 1, all of the chemical and biochemical reactions preceding the molecular initiating event (upstream events) require chemical speciation models and metabolic simulators to forecast the relative probabilities of occurrence for important chemical species. Predicting the possible interactions leading up to the molecular initiating events is the role of QSAR modelling.

The sequence of biological effects leading from the molecular initiating events to a specific adverse outcome (downstream events) is a separate portion of the overall toxicity pathway and requires biological models to predict the ultimate consequence of the molecular initiating event. These biological models may often be the same models used in many *in vitro/in vivo* endpoint extrapolations; however, the proposed framework offers an important advantage for the development of endpoint extrapolation models. *In vitro/in vivo* extrapolations may have a model application domain that is not readily apparent from empirical test data alone. Mixing chemicals from different domains in an *in vitro/in vivo* endpoint extrapolation model can cause the same prediction errors as would be found if QSAR models were used outside the model domain. In the proposed conceptual framework, information from the domain of application for a chemical can be used to confine endpoint extrapolation models to more mechanistic domains and make them more reliable.



As a central organising principle, therefore, the International QSAR Foundation is identifying important molecular initiating events involving binding of chemicals with important cellular targets. From that starting point, each specific initiating event is used to group chemicals based on the mechanistic ability of the chemical to cause the molecular initiating event to occur. From those same molecular initiating events, the biological effects that result from the disturbance are grouped by the level of biological organisation represented by the effect endpoints as well as the route of exposure and dose regime. Since a given reactive toxicant may interact with numerous cellular targets, thereby causing numerous biological effects, the potential of a chemical to cause a specific biological effect must be estimated by understanding the relative reactivity, or selectivity, of the chemical toward an array of cellular targets. A systematic description of the selectivity of reactive chemicals for cellular targets is the reactivity profile of chemicals (Shultz et al., 2005).

This framework improves the endpoints that are to be modelled by QSAR by removing all of the variability introduced by the biological testing. In short, it increases the separation of the relative problems of predicting chemical reactions with QSAR models and the biological consequences with biological models. The approach is to begin generating systematic databases for electrophile/nucleophile reactivity across the spectrum of soft to hard electrophile/nucleophiles. These data are used, in turn, as independent variables to explore for correlations with downstream effects. When potential models are found, the training sets are redesigned to test the causal relationships.

References

Bradbury, S. P., Feijtel, T. C. J., and van Leeuwen, C. J. (2004). Meeting the scientific needs of ecological risk assessment in the regulatory context. *Environmental Science &Technology, 465A*, December 1.

Jaworska, J. S., Comber, M., Auer, C., van Leeuwen, C. J. (2003) Summary of a workshop on regulatory acceptance of (Q)SARs for human health and environmental endpoints. *Environmental Health Perspectives 111*, 1358-1360.

Shultz, T. W., Carlson, R. E., Cronin, M. D. T. et al. (2005). A conceptual framework for predicting toxicity of reactive chemicals: Models for soft electrophility. *Proceedings of the Knoxville Reactivity Workshop*, May 8-11.

Correspondence to

Gilman D. Veith
International QSAR Foundation to Reduce Animal Testing
Two Harbors, Minnesota 55616
USA
e-mail: GDVeith@earthlink.net

Session 6.2
Biokinetic modelling *in silico*

Integration of PBPK and Reaction Network Modelling: Predictive Xenobiotic Metabolomics

Raymond S. H. Yang[1], Arthur N. Mayeno[1], Kai H. Liao[1,2], Kenneth F. Reardon[1,2] and Brad Reisfeld[1,2]
[1]Quantitative and Computational Toxicology Group, [1]Department of Environmental and Radiological Health Sciences and
[2]Department of Chemical and Biological Engineering, Colorado State University, Fort Collins, CO, USA

Summary
Our research group has aimed to integrate computational modelling with in vitro and in vivo experimentation towards the advancement of chemical mixture toxicology while minimising animal use. In the case of complex chemical mixtures and their interactions, the computer-assisted approach of Biochemical Reaction Network Modelling offers a ray of hope. The possible linkage between this novel computational methodology and physiologically-based pharmacokinetic (PBPK) modelling could result in a multi-scale computer simulation platform capable of predicting complex pathway interactions and metabolite concentrations at the molecular level up to tissue and organ concentrations at the whole organism level.

Keywords: PBPK modelling, biochemical reaction network modelling

Introduction

Computer simulation has been used extensively in the physical sciences and engineering and its wide applications touch every aspect of our lives. One of the best examples to illustrate the maturity and utility of computer simulation is the modern jetliner. The Boeing 777 is the first jetliner to be 100 percent digitally designed using three-dimensional CAD/CAM (computer-aided design/computer-aided manufacturing) technology (http://www.boeing.com/commercial/777family/background.html). Throughout the design process, the airplane was "pre-assembled" on the computer, eliminating the need for a costly, full-scale mock-up. In contrast, the utilisation of computational technology in biology lags far behind that in physical sciences and engineering. This deficiency, as well as the importance of integrating computer technology into biology, has been recognised by other investigators, as reflected by the following quote from Craig Venter (Butler, 1999): "…If we hope to understand biology, instead of looking at one little protein at a time, which is not how biology works, we will need to understand the integration of thousands of proteins in a dynamically changing environment. A computer will be the biologist's number one tool…"

Computer simulation of biological processes is a realistic and workable method to either replace animal studies, or at least, reduce animal usage in experimentation. Ethical reasons aside, the reality is that experimental toxicology cannot keep pace with the number of chemicals in commerce (about 80,000), plus the new synthetic organic chemicals coming on stream (about 2,000/year). For instance, the U.S. National Toxicology Program with its predecessor, the National Cancer Institute, has conducted chronic toxicity/carcinogenicity studies on about 600 chemicals in their combined effort over the last 43 years. This represents a minuscule portion of the chemicals that are being used today. It is quite clear that we will never be able to catch up under the present mode of operation. Considering further the potential presence of numerous chemical mixtures, it is certain that alternative methods, such as computational tools, must be developed to handle the backlog more efficiently.

Our laboratory has been working on a possible solution to the above dilemma. We advocate the integration of physiologically

based pharmacokinetic (PBPK) modelling with biochemical reaction network (BRN) modelling to create multi-scale computer simulations. These models would provide predictions of the fate of a chemical or chemical mixture from the level of the whole organism down to molecular interactions (Klein et al., 2002; Liao et al., 2002; Reisfeld and Yang, 2004; Yang et al., 2004a; Mayeno et al., 2005). This methodology is currently at the research and development stage. When completed, chemicals or chemical mixtures with little or no animal toxicity data can be fed into the computer simulation program and their potential adverse health effects deduced from the metabolic reaction networks generated. In the sections below we will first discuss PBPK and BRN modelling, then the integration of these two types of models, and finally the concept of predictive xenobiotic metabolomics.

Materials and methods

PBPK modelling

PBPK is a special type of pharmacokinetics where physiology and anatomy of the animal or human body, and the biochemistry of the chemical or chemicals of interest, are incorporated into a conceptual model for computer simulation. Unlike classical pharmacokinetics, PBPK modelling is a powerful tool for many types of extrapolation: inter-species, inter-routes, inter-doses, inter-life stages, etc.

The concept of PBPK had its embryonic development in the 1920s. PBPK modelling blossomed and flourished in the late 1960s and early 1970s in the chemotherapeutic area due mainly to the efforts of investigators with expertise in chemical engineering. In the mid 1980s, work on PBPK modelling of volatile solvents started yet another "revolution" in the toxicology and risk assessment arena. Today, there are more than 1000 publications directly related to PBPK modelling of industrial chemicals, drugs, environmental pollutants, and simple and complex chemical mixtures. A book on PBPK modelling has recently been published from our laboratory in collaboration with others (Reddy et al., 2005).

The fundamentals of PBPK modelling are to identify the principal organs or tissues involved in the disposition of the chemical of interest and to correlate the chemical absorption, distribution, metabolism, and excretion (ADME) within and among these organs and tissues in an integrated and biologically plausible manner. A scheme is usually formed where the normal physiology is followed in a graphical manner (i.e., a conceptual model). Within the boundary of the identified compartment (e.g., an organ or tissue), whatever "comes in" must be accounted for via whatever "goes out" or whatever is transformed into something else. This mass balance is expressed as a mathematical equation with appropriate parameters carrying biological significance. A series of such mass balance equations representing all of the interlinked compartments is formulated to express a mathematical representation, or model, of the biological system. This model can then be used for computer simulation to predict the time course behaviour of any given parameter.

For more detailed information on PBPK modelling and its related methodologies, readers are referred to two recent publications (Yang et al., 2004b; Reddy et al., 2005).

Biochemical reaction network modelling

Biochemical reaction network (BRN) modelling has its origin in chemical and petroleum engineering. It was successfully employed in computer modelling and simulation of the complicated processes in oil refineries. A reaction network model is a tool that is used to predict the amounts of reactants, intermediates, and products as a function of time for a series of coupled chemical reactions (potentially numbering in the tens of thousands of reactions). Broadbelt et al. (1994, 1996) refined previous ideas and used concepts of graph theory to represent species connectivity. They also made use of computational quantum chemistry and linear free energy relationships (LFER) to automate the process of determining reaction rate constants. The essential idea is that the model takes, as input, specifications for the reactants (usually in terms of their chemical structures), as well as rules stipulating the nature of the relevant chemical reactions. Algorithms within the reaction network model develop the associations between species and create and solve the controlling kinetic equations in the reaction model. Thus, the outputs of the simulation are the connections between reaction species as well as the concentrations of these species over time. The idea of using reaction network modelling for biomedical applications was put forth by a joint effort between our laboratory and Rutgers University (Klein et al., 2002).

Over the last four years, a programme, BioTRaNS (**Bio**chemical **T**ool for **R**eaction **N**etwork **S**imulation), was created *de novo* in our laboratory based on an extensive review of the literature on molecular modelling of substrates of P450 enzyme systems. BioTRaNS integrates modules of our own creation, as well as existing software and database tools, such as CORINA (molecule structure prediction), MOPAC7 (quantum chemical calculations), GraphViz (mathematical graph visualisation), and Daylight and OpenEye toolkits (symbolic molecule manipulation and chemical/biochemical reaction transformations and prediction). As part of the BioTRaNS effort, other methods (quantitative structure activity relationships, decision trees) have also been developed for the prediction of the probabilities of cytochrome P450 binding of chemicals.

The novelty and advantages of BioTRaNS over previous and other frameworks are as follows:

a) It was written from the ground up to focus on biological applications rather than on petrochemical applications.
b) It allows user-friendly programme usage and interaction.
c) It specifically considers enzyme-substrate interactions.
d) It has "hooks" or interfaces to communicate and interact with PBPK or other modelling tools.
e) It uses a cheminformatics-industry standard means of representing molecules.
f) It is flexible and user-friendly in terms of specifying reactions and reaction feasibility.
g) It uses a dynamically-updated database of molecules and molecular properties.
h) It is well documented for the user.
i) It has a well-defined and documented application programming interface (API) for programmers to use when designing applications to interact with the present application.

A simplified description and information flow is illustrated in figure 1 (see Mayeno et al., 2005 for a more detailed description). In brief, BioTRaNS takes a description of a set of chemicals and the enzymes believed to be involved in their metabolism and produces the detailed metabolic pathways, showing the interconnections between the metabolites as well as the concentrations of all of these chemical species over time.

a) First, the user inputs the concentration of a single chemical (or concentrations of individual components in a chemical mixture) as well as the types and amounts of enzymes that the user selects to act on the chemicals.

b) The reactants, along with data from the Molecule Property Database, are fed into the Feasibility Module, which computes the probability that each given reactant will be a substrate for each of the specified enzymes. If the probability of an interaction is below a user-defined threshold, the reactant does not undergo chemical transformation via that enzyme. Even though our initial development of BRN modelling has specifically focused on certain types of chemical mixtures, this feasibility module will be further developed to give the software the capability to predict substrate feasibility of any chemical that will be studied in the future.

c) The Transformation Module, with chemical transformation information from the Reaction Rules Database, performs the "virtual chemistry" of transforming the reactants to products. These products in turn become reactants that are checked for reaction feasibility.

d) The Pathway Module uses the metabolites generated from step "c" to establish the connections between all metabolites, forming the basis for the reaction network. The network structure or topology can be used to give various "views" of the reaction interconnections.

e) Simultaneous to pathway creation, the Kinetics Module creates the reaction rate equations (ordinary differential equations) and retrieves the reaction rate constants for these reactions from the Rate Constant Database, or estimates them from related data, such as physicochemical, electronic, and quantum mechanical properties.

f) The total set of rate equations is then solved by the Solver Module to determine the concentrations of all species as a function of time.

From the information set produced by BioTRaNS an investigator can examine the nature and lifetimes of species of interest and, in the context of health risks, easily locate highly reactive species. Moreover, due to its design, information can be fed back and forth between BioTRaNS and lower level (e.g., enzyme induction) and higher level (organ/organism level) modelling tools to give a more complete picture of the risk.

Results

Biochemical reaction network modelling of a mixture of four volatile organic chemicals

As an example of the use of BRN modelling to examine interactions, we will use a mixture of four volatile organic chemicals (VOCs): trichloroethylene (TCE), chloroform (CF), tetrachloroethylene (Perc), and 1,1,1-trichloroethane (MC). All four

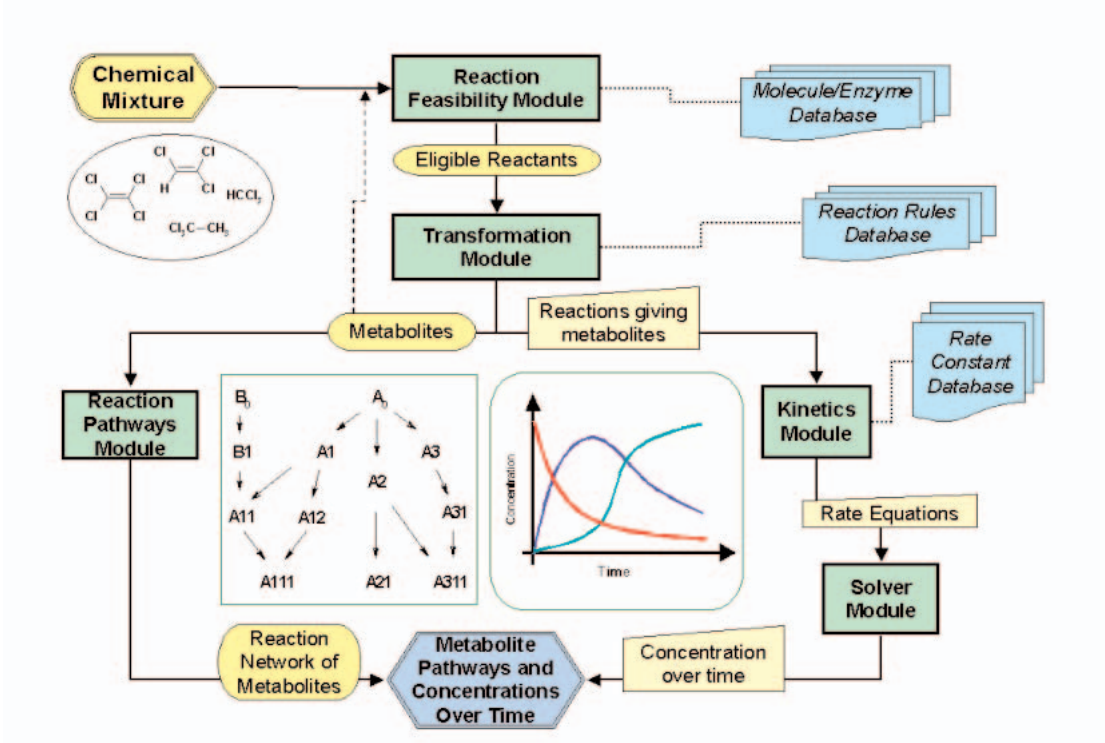


Fig. 1: Information flow through the BioTRaNS

are prevalent drinking water or ground water contaminants and they are likely to be present in such media together. This study was reported in a recent publication (Mayeno et al., 2005).

The interconnected biotransformation pathways of all four VOCs (fig. 2) illustrate the close relationship among the metabolic pathways of these chemicals, their shared enzyme systems, the potential for generating the same reactive species from different parent compounds, and the dynamic interactions among the linked pathways influencing the possible outcome of toxicities. Our laboratory has completed the qualitative aspects of the biochemical reaction network modelling of the four volatile organic chemicals based on biochemical reaction mechanisms of the relevant CYP and related enzymes. Moreover, we have incorporated enzyme-reaction mechanisms to help predict metabolite formation. For instance, the mechanism-based biotransformation of TCE, as generated by BioTRaNS, is shown in figure 3. The first step of CYP2E1-catalysed oxidation involves the formation of an intermediate between the high-valent iron-oxo complex, $[(FeO)^{3+}]$ of the CYP haeme and the alkene, as postulated by Miller and Guengerich (1982). This intermediate has been linked mechanistically to the 1,2-shifts of Cl (or H), leading to the formation of an aldehyde (or acid chloride), as shown in figure 3. Details of these (bio)chemical processes and subsequent step-wise reaction mechanisms, with their respective

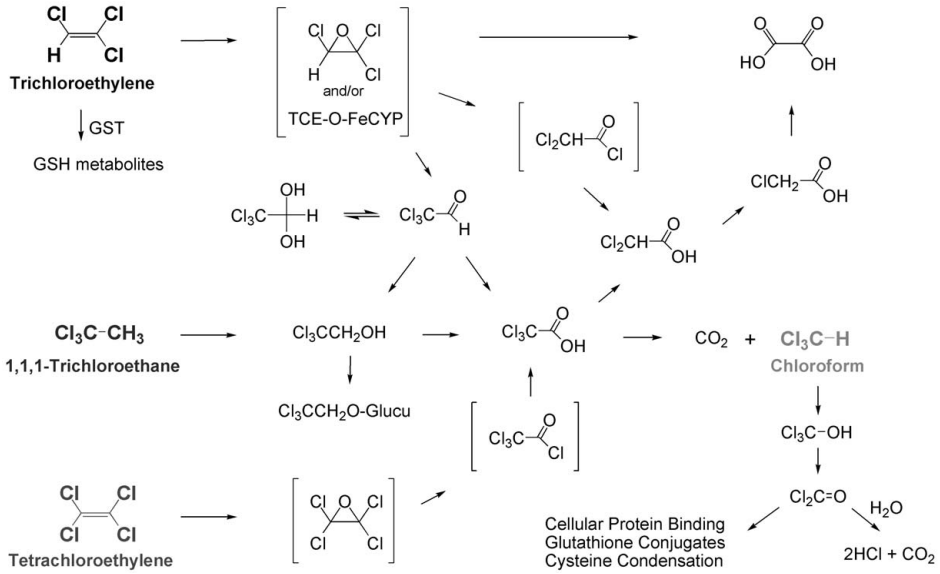


Fig. 2: Combined metabolic pathways of trichloroethylene, chloroform, tetracholoroethylene, and 1,1,1-trichloroethane (only major pathways shown)

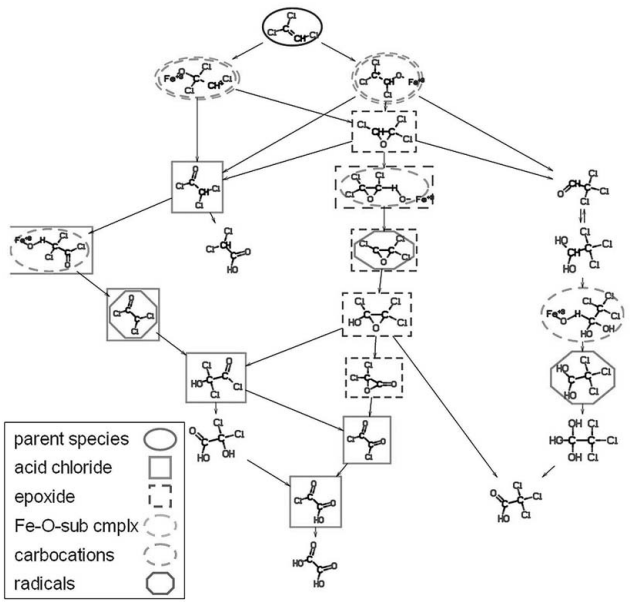


Fig. 3: A portion of the BioTRaNS-generated biotransformation pathways using a postulated mechanism for CYP-mediated TCE oxidation (Miller and Guengerich, 1982). Reactive metabolites are highlighted as follows: epoxides (brown, box, dashed); acid chlorides (orange, box, solid); and starting chemicals (blue, ellipse, solid). Please see website www.altex.ch for figure 3 in colour. (Reprinted with permission from Mayeno et al., 2005. Copyright (2005) American Chemical Society).

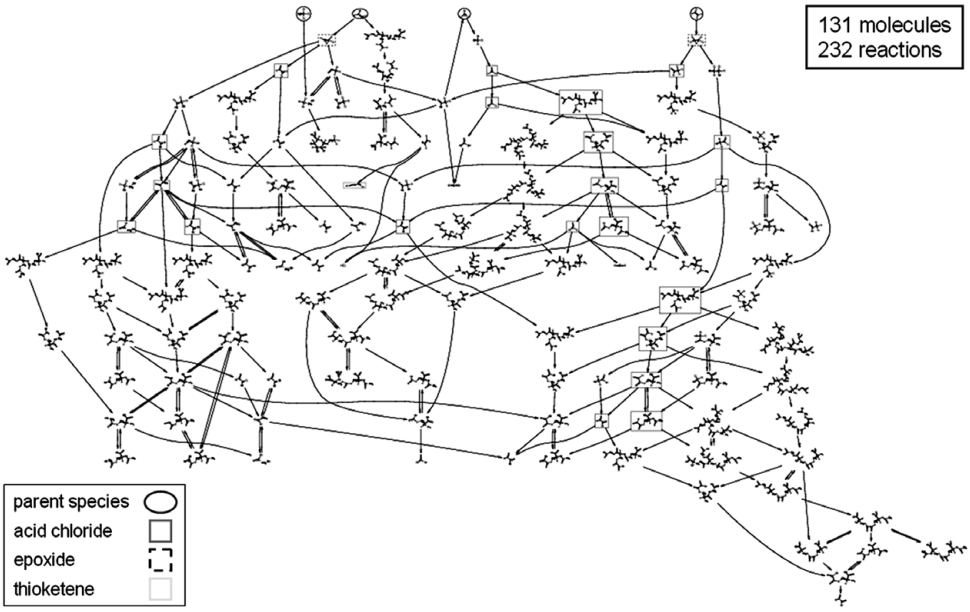


Fig. 4: BioTRaNS-generated biotransformation reaction network for four volatile organic chemicals: trichloroethylene, tetrachloroethylene, methyl chloroform, and chloroform. The software generated this figure based on reaction rules and interconnected the pathways via metabolites in common. Reactive species were highlighted in red boxes, after substructures (SMARTS) of these species were input by the user. Please see website www.altex.ch for figure 4 in colour. (Reprinted with permission from Mayeno et al. 2005. Copyright (2005) American Chemical Society)

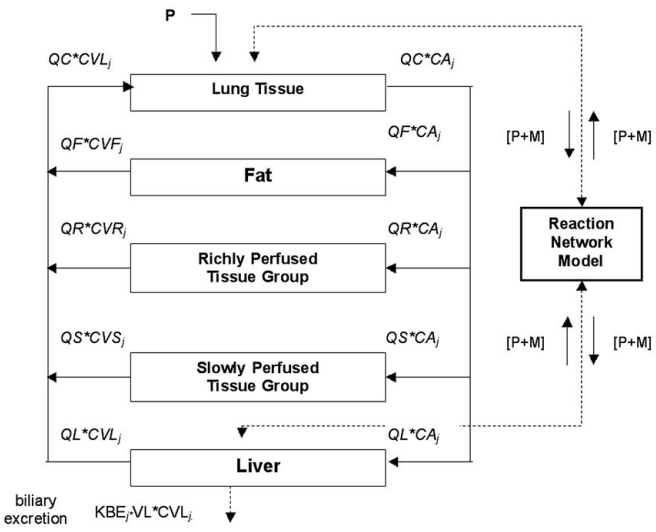


Fig. 5: A graphical representation of a preliminary conceptual integrated physiologically-based pharmacokinetic/biochemical reaction network model for benzo[a]pyrene.
P = parent compound (BaP); M = metabolites; Q = flow rate; CA = arterial blood; CV = venous blood.

originally published sources, have been described in our recent paper (Mayeno et al., 2005). Our computer-generated interconnected metabolic pathways for these four VOCs (fig. 4) demonstrate the metabolite inventory and interconnections ("predictive xenobiotic metabolomics"). The predicted pathways (fig. 4) match the known metabolic pathways (fig. 2) very well.

Linking physiologically-based pharmacokinetic models and biochemical reaction network models

Since PBPK and BRN models address different aspects of predictive toxicology, linking them will yield a powerful tool. To this end, the development and validation of an integrated BRN/PBPK model of benzo[a]pyrene (BaP) was studied (Liao, 2004). In this case, the BRN model used was a predecessor to the BioTRaNS model described earlier. figure 5 illustrates the interconnection between the PBPK and BRN models. The linkage between the two models, in this case, occurs in the liver compartment (the major organ responsible for metabolism) via the venous circulation.

The PBPK model predicts the ADME of BaP and metabolite(s) of interest circulated to the organs, while the BRN model calculates the amounts and rates of metabolites formed and unmetabolised parent chemical(s) that can be distributed back to organs/tissues via PBPK modelling. A general equation, for any tissue or organ, is:

$$Vi\frac{dC_{ij}}{dt} = Q_i\left(CA_j - CV_{ij}\right) - \mathrm{Metab}_{ij} - \mathrm{Elim}_{ij} + \mathrm{Absorp}_{ij} - \mathrm{Pr}\,Binding_{ij}$$

where V_i represents the volume of tissue group *i*, Q_i is the blood flow rate to tissue group *i*, CA_j is the concentration of chemical *j* in arterial blood, and C_{ij} and CV_{ij} are the concentrations of

chemical *j* in tissue group *i* and in the effluent venous blood from tissue *i*, respectively. $Metab_{ij}$ is the rate of metabolism for chemical *j* in tissue group *i*; it can be predicted by BRN models in the liver and other metabolising organs and is equal to zero in other tissue groups. $Elim_{ij}$ represents the rate of elimination from tissue group *i* (e.g., biliary excretion from the liver), $Absorp_{ij}$ represents uptake of the chemical from dosing (e.g., oral dosing), and $PrBinding_{ij}$ represents protein binding of the chemical in the tissue.

Figures 6 and 7 are representative simulations of the integrated model against published data in the literature. The high quality of the simulations, which were obtained by specifying reaction rules, is remarkable. It is noteworthy that our BRN modelling indicated that the biotransformation of BaP, a single compound, involves 246 possible reactions and 150 possible metabolic products (Liao, 2004). These results demonstrate the potential for complexity, the capability of BRN modelling, and the promise of linked PBPK-BRN models.

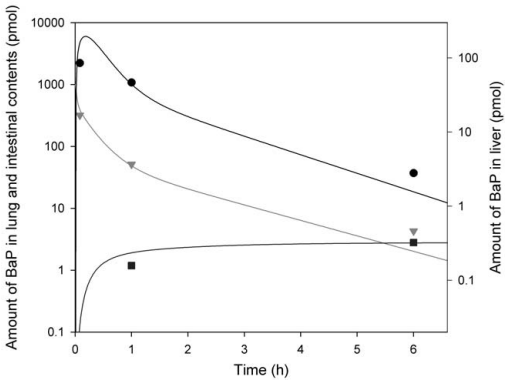


Fig. 6: BaP in liver (●), lung (▲), and intestinal contents (■) of rats after exposure to BaP intra-tracheally at 1 mg/kg body weight. (Data from Weyand and Bevan, 1986)
Solid lines represent biochemical reaction network (BRN)/ physiologically-based pharmacokinetic (PBPK) model simulations.

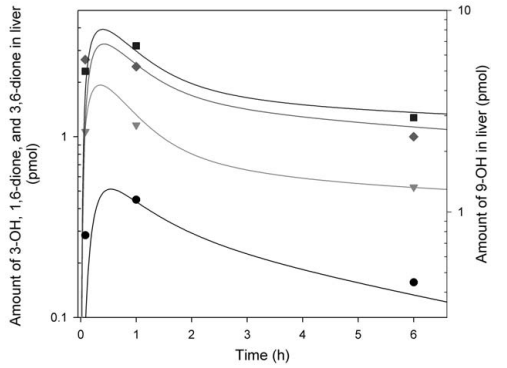


Fig. 7: Amount of BaP metabolites, 3-OH (▼), 9-OH (●), 1,6-dione (◆), and 3,6-dione (■), in liver of rats after exposure to BaP intra-tracheally at 1 mg/kg body weight. (Data from Weyand and Bevan, 1986)
Solid lines represent BRN/PBPK model simulations.

Discussion and conclusions

Metabolomics is the study of the metabolites contained in a human or animal cell, tissue or organ and involved in primary and intermediary metabolism. It is an emerging "–omics" technology that has already shown great promise in providing important and relevant health-related information (Brindle et al., 2002).

Since xenobiotics, including drugs, affect our health in both positive and negative ways, it is reasonable to expand the domain of metabolomics to include xenobiotic reaction network pathways in our assessment of human health. With the linkage of PBPK and BRN modelling as described above and the possible linkage to other modelling tools (fig. 8), the field of "*Predictive Xenobiotic Metabolomics*" can be advanced. Using such an approach towards chemical and chemical mixture toxicology, reduction and even total avoidance of animal usage can be realised in the future.

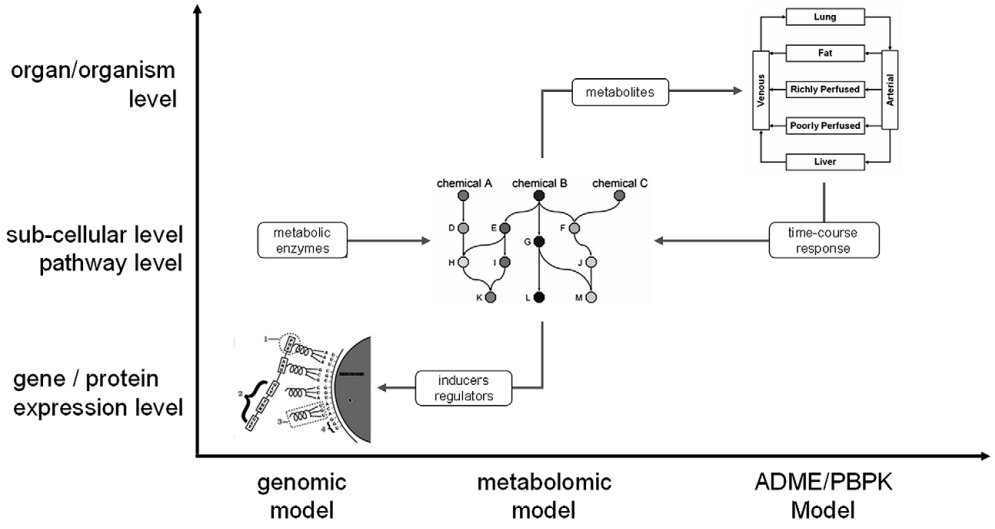


Fig. 8. Multi-scale modelling and systems biology approach towards global toxicological effects of a chemical or chemical mixtures.

References

Brindle, J. T., Antti, H., Holmes, E. et al. (2002). Rapid and non-invasive diagnosis of the presence and severity of coronary heart disease using 1H-NMR-based metabonomics. *Nature Medicine 8*, 1439-1444.

Broadbelt, L. J., Stark, S. M. and Klein, M. T. (1994). Computer generated reaction networks: on-the-fly calculation of species properties using computational quantum chemistry. *Chem. Eng .Sci. 49*, 4991-5010.

Broadbelt, L. J., Stark, S. M. and Klein, M. T. 1996. Computer generated reaction modelling: decomposition and encoding algorithms for determining species uniqueness. *Comp. Chem. Eng. 20*, 113-129.

Butler, D. (1999). Computing 2010 from black holes to biology. *Nature 402*, C67-C70.

Klein, M. T., Hou, G., Quann, R. et al. (2002). BioMOL: A computer-assisted biological modelling tool for complex chemical mixtures and biological processes at the molecular level. *Environ. Health Perspect. 110 (Supplement 6)*, 1025-1029.

Liao, K. H., Dobrev, I., Dennison, Jr., J. E. et al. (2002). Application of biologically based computer modelling to simple or complex mixtures. *Environ. Health Perspect. 110 (Supplement 6)*, 957-963.

Liao, K. H. (2004). *Development and Validation of a Hybrid Reaction Network/ Physiologically Based Pharmacokinetic Model of Benzo[a]pyrene and Its Metabolites*, Ph.D. Dissertation, Department of Chemical Engineering, Colorado State University, 214 pp.

Mayeno, A. N., Yang, R. S. H. and Reisfeld, B. (2005). Biochemical Reaction Network Modelling: A New Tool for Predicting Metabolism of Chemical Mixtures. *Environ. Sci. Tech. 39,* 5363-5371.

Miller, R. E. and Guengerich, F. P. (1982). Oxidation of Trichloroethylene by Liver Microsomal Cytochrome P-450: Evidence for Chlorine Migration in a Transition State Not Involving Trichloroethylene Oxide. *Biochemistry 21*, 1090-1097.

Reddy, M., Yang, R. S. H., Clewell III, H. J. and Andersen, M. E. (2005). *Physiologically Based Pharmacokinetics: Science and Applications*, 420 pp. Hoboken, NJ: John Wiley & Sons, Inc.

Reisfeld, B. and Yang, R. S. H. (2004). A reaction network model for CYP2E1-mediated metabolism of toxicant mixtures. *Environ. Toxicol. Pharmacol. 18*, 173-179.

Yang, R. S. H., El-Masri, H. A., Thomas, R. S. et al. (2004a). Chemical mixture toxicology: from descriptive to mechanistic, and going on to in silico toxicology. *Environ. Toxicol. Pharmacol. 18*, 65-81.

Yang, R. S. H., Andersen, M. E., Dennison, J. E. et al. (2004b). Physiologically Based Pharmacokinetic and Pharmacodynamic Modelling. In Ed. E. C. Holland (ed.), *"Mouse Models of Cancer"* (pp. 391-405). New York, NY: Wiley Inc.

Acknowledgements

This research could not have been possible without the generous support of NIEHS (Superfund Basic Research Program Project P42 ES05949; R01 ES09655; T32 ES 07321; K25 ES11146; K25 ES012909), ATSDR/CDC (U61/ATU881475; RO1 OH07556), and US Air Force (F33615-91-C-0538; F49620-94-1-0304).

Correspondence to

Dr. Raymond S. H. Yang
Department of Environmental and Radiological Health Sciences
Colorado State University
1680 Campus Delivery
Fort Collins, CO 80523, USA
e-mail: raymond.yang@colostate.edu

Session 6.3
Computational toxicology

Comparative Analysis of Gene Networks at Multiple Doses and Time Points in Livers of Rats Exposed to Acetaminophen

Hiroyoshi Toyoshiba[1], Hideko Sone[1], Fred Parham[2], Richard D. Irwin[2], Gary A. Boorman[2] and Christopher J. Portier[2]

[1] National Institute for Environmental Studies, Tsukuba, Ibaraki, Japan; [2]National Institute of Environmental Health Sciences, RTP, NC, USA

Summary

Gene interaction networks (GINs) have been developed as a potential tool to understand biological mechanisms or responses to chemical toxicity. Our recent studies developed two algorithms to identify and quantify gene interaction networks. One is to identify networks based on a statistical criteria and the other is to test hypothetical networks using microarray data. These methods were combined with clustering analysis, networks were clustered and father analysis was performed to quantify linkages in the networks and to follow microarray change over time and dose levels in the model of rat exposed to acetaminophen. These analysis methods could be used for risk assessment and could improve the mechanistic understanding of chemical toxicity.

Keywords: gene interaction networks, computational biology, risk assessment, microarray

Introduction and questions

Several methods of microarray data analysis have been developed in order to gain a mechanistic understanding of chemical toxicity. Clustering analysis (Bharadwaj et al., 2005; Hamadeh et al., 2002; Pan et al., 2003) has become a powerful tool to characterise chemical toxicity through the patterns of expression levels of thousands of genes. Recently, a number of approaches have tried to understand the mechanisms of chemical toxicity by means of pathway analysis or linkage analysis between genes. Gene interaction networks (GINs) have been developed to analyse links between genes to understand the biological mechanisms or responses to chemical toxicity. Several methods have been developed under different mathematical frameworks, i.e. ordinary differential equations (Chen et al., 1999; Moles et al., 2003) and Boolean networks etc. Among these methods, the Bayesian networks (Friedman et al., 2000) approach is a powerful method by which acyclic directed networks are used to characterise networks. We have recently developed two algorithms based on statistical criteria and statistical quantification of linkages in GINs. Both methods were developed based on Bayesian networks: the TAO-Gen (Theoretical Algorithm for identifying Optimal Gene interaction networks) algorithm (Yamanaka et al., 2004) is used to identify the best GIN based on a statistical criteria, and a statistical algorithm using log-linear function form (Toyoshiba et al., 2004) is used to quantify linkages between genes. In this manuscript, father analysis of the network structure was introduced by applying the method to data measured from livers of rats 6, 24 and 48 hours after exposure to 50 mg/kg, 150 mg/kg and 1500 mg/kg of acetaminophen. These nine data sets were acquired with 10 replicates. For each of the data sets, the best network was estimated by TAO-Gen algorithm. The networks were clustered into two groups by k-mean clustering. The clustering analysis separated the networks into the high dose group (1500

mg/kg) and low dose group (50 mg/kg and 150 mg/kg). Using our method (Toyoshiba et al., 2004), each linkage in the networks was quantitatively analysed to identify whether the linkage was an activator or inhibitor. The results of the analysis showed that the network may change with dose level and that the relationship of the linkage could change over the different time points.

This method offers a new approach to analyse GINs quantitatively and is capable of improving dose-response analysis and risk assessments. Also, the method developed here could contribute to reduction of the use of animals in experiments in the future. For example, the method showed the clustering networks and combination of network analysis and clustering analysis could provide pattern recognition as well as linkage information. Chemicals with the same network structure could be predicted to have the same spectrum of effects on humans and animals. Hence, studies on a specific chemical could be used to predict the toxicity of other chemicals with the same network structures.

Materials and methods

Estimation procedure

Let us define the GINs. $G = [G_1, G_2, \ldots G_P]$ represents P random variables associated with P genes with g_i (i=1,2,…P) as the actual observation of G_i. Let g_i be $nx1$ matrix denoting actual n samples from gene G_i. Log-linear functional form is defined as the following,

$$\log(g_j) = \sum_{i=1}^{P} I_{ij} \cdot \beta_{ij} \cdot \log(g_i) + \varepsilon_j \qquad (1)$$

where I_{ij} represents an indicator function having 1 if there exists a linkage from gene G_i to gene G_j and 0 if there is no linkage. β_{ij} is a coefficient in the log-linear regression and finally, ε_j is assumed to follow the normal distribution $N(0, \sigma_j^2)$. Hence, the model has I_{ij}, β_{ij} and σ_j^2 as the parameters and the first two parameters are particularly interesting. The indicator functions can determine the network structure and the coefficient β_{ij} is obtained by the mean of the posterior distribution which tells whether the linkage is statistically significant (Toyoshiba et al., 2004). A matrix consisting of I_{ij} called transition matrix determines the structure of the network and a matrix consisting of β_{ij} is called parameter matrix. The transition matrix is estimated by TAO-Gen algorithm (Yamanaka et al., 2004). Once the transition matrix is given, the parameter matrix can be estimated with the posterior distribution of β_{ij} (Toyoshiba et al., 2004). The method tells whether the linkage is an activator or inhibitor and also the level of significance. If a gene has a regulatory effect on G_i, then that gene is referred to as a "Parent of G_i" and we refer to it as belonging to the set Pa(Gi).

Clustering networks

The transition matrix can be considered to be the single sequence consisting of two elements 0 and 1. The last element of the i^{th} row precedes the first element of the $i+1^{th}$ row. Hence, the network structure can be summarised as a single sequence. Clustering analysis can be performed with these sequences to elucidate similarities or differences between networks. The distance between two sequences was defined as the sum of the absolute deference of each element of the two sequences. K-means clustering was performed and the networks were clustered into two groups and two centred networks were obtained.

Quantification of the centred networks

After estimating two centred networks, each path in the centred networks was quantified by Markov Chain Monte Carlo sam-

Tab. 1: Seventeen (17) genes used to develop GINs to understand how the apoptosis and oxidative stress genes are linked after exposed to Acetaminophen (APAP).

	Abbreviation	Clone ID	Description
Apoptosis	MADH2	AA818109	*Rattus norvegicus* Smad2 protein (Smad2) mRNA, complete cds
	NFKB1	AA858801	*Rattus norvegicus* nuclear factor kappa B p105 subunit mRNA, 3' end
	AKT1	AA858883	Rat mRNA for RAC protein kinase alpha, complete cds
	TP53	AA875052	Rat mRNA for nuclear oncoprotein p53
	PCNA	AA924358	Rat proliferating cell nuclear antigen (PCNA/cyclin) mRNA, complete cds
	SRC	AA926101	ESTs, Highly similar to NEURONAL PROTO-ONCOGENE TYROSINE-PROTEIN KINASE SRC (*M.musculus*)
	PPARG	AI111890	*Rattus norvegicus* mRNA for PPAR-gamma protein, complete cds
Oxidative stress	MAPK3	AA875555	*Rattus norvegicus* microtubule-associated protein-2 kinase mRNA, partial cds
	GSK3B	AA964758	*R. norvegicus* mRNA for tau protein kinase I
	MAPK1	AI058662	ESTs, Highly similar to Phosphorylated Map Kinase Erk2 (*R.norvegicus*)
	CDKN1B	AI071529	*Rattus norvegicus* mRNA for p27, complete cds
	GRB2	AI071530	Rat mRNA for Ash-m, complete cds
	NFE2L2	AI071579	*Rattus norvegicus* NF-E2-related factor 2 mRNA, complete cds
	ITGB1	AA819911	*Rattus norvegicus* integrin beta-1 subunit (integrin beta-1) mRNA, complete cds
Significant changed genes	CASP2	AA956222	*Rattus norvegicus* caspase 6 (Mch2) mRNA, complete cds
	MT1a	AA900218	Rat metallothionein-i (mt-1) mrna
	HMOX1	AA874884	Rat heme oxygenase gene, complete cds

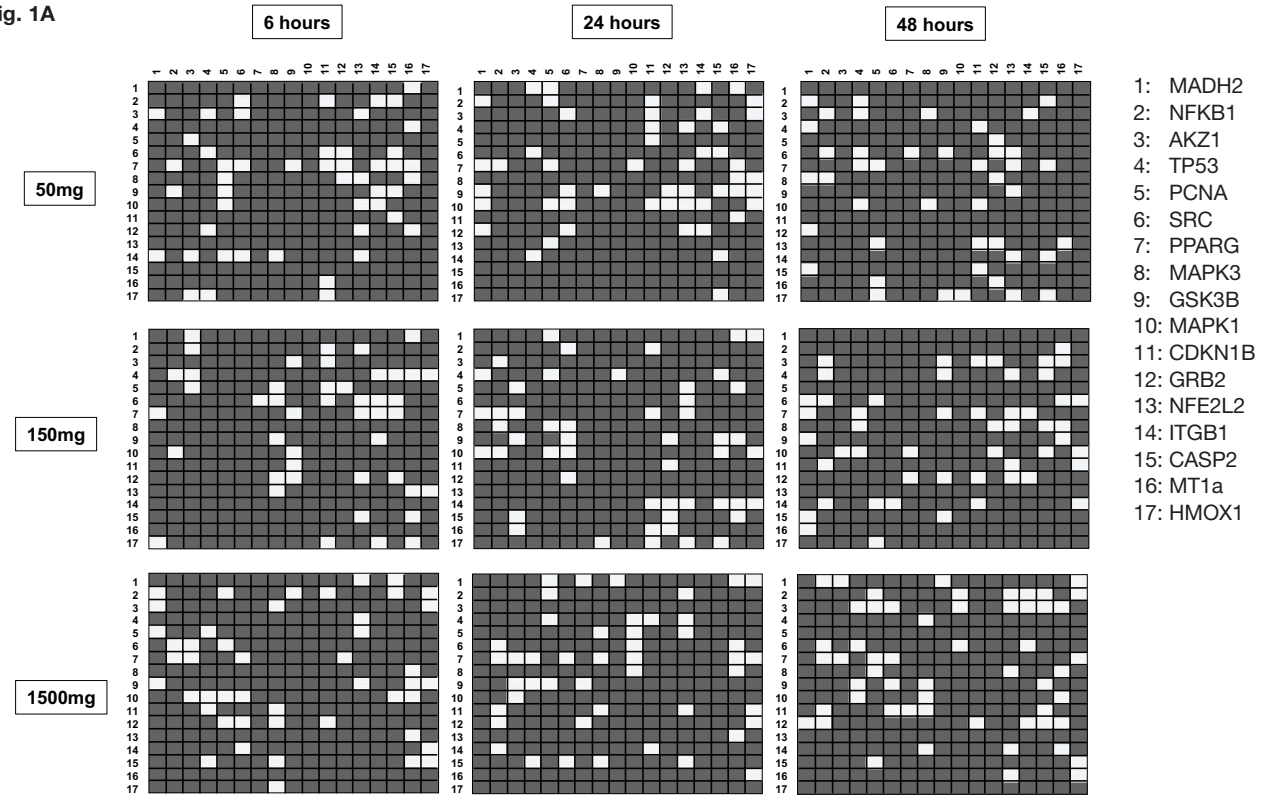
pling methods as described (Toyoshiba et al., 2004). The quantification defines the statistical significance of the linkages between the different genes and the quality of each linkage (activator or inhibitor).

Results

Seventeen genes (MADH2, NFKB1, AKT1, TP53, PCNA, SRC, PPARG, MAPK3, GSK3B, MAPK1, CDKN1B, GRB2, NFE2L2, ITGB1, CASP2, MT1a, HMOX1) from a single gavage dose study in Fischer rats at multiple doses and multiple times after exposure to acetaminophen were identified as potentially belonging to a network governing the liver response (tab. 1).

The genes can be divided into three basic groupings; apoptosis (Apt) related genes (MADH2, NFKB1, AKT1, TP53, PCNA, SRC, PPARG), oxidative stress (Ost) related genes (MAPK3, GSK3B, MAPK1, CDKN1B, GRB2, NFE2L2, ITGB1) and other (Oth) significantly elevated genes (CASP2, MT1a, HMOX1). CASP2 was selected because of its known role in stimulating apoptosis (Harvey et al., 1997), activating HMOX1 (Panahian et al., 1999) and protecting against oxidative stress (Beattie et al., 2005). The TAO-Gen algorithm was applied to the data for each dose (50, 150 and 1500 mg/kg) at each follow-up time (6, 24 and 48 hours) resulting in nine separate models (fig. 1A). Significance for each linkage in the TAO-Gen algorithm was assessed using Markov Chain Monte Carlo and only the significant linkages were retained. Figure 1A is a graphical representation of these linkages. In this representation using the matrix form, paths lead from genes in columns to genes in rows. For example, in the graph of 50 mg/kg at 6 hours, AKT1 in the third row has paths to MADH2, TP53, SRC and NFE2L2. These linkages are highlighted in yellow in the 1st, 4th, 6th and 13th columns of the third row. To compare these networks, k-means clustering (k=2) was applied to nine transition matrices to determine which are closer structures. The resulting two clusters split the networks by dose and not by time (fig. 2). The first cluster contained all times for the low (50 mg/kg) and middle (150 mg/kg) dose groups, whereas the other cluster contained all times for the high dose (1500 mg/kg) group. The two centred

Fig. 1A



$$\log(g_j) = \sum_{i=1} t_{ij} \cdot p_{ij} \cdot \log(g_i) + \varepsilon_j$$

Fig. 1B

Fig. 1A: Nine estimated GINs by dose and exposure time in male rats exposed to acetaminophen where yellow boxes indicate linkages between genes and green indicates no linkage with the parent gene on the vertical axis and the child on the horizontal axis.

Fig. 1B: Linkages from AKT1 to other genes in the low dose groups at 50, 150 and 1500 mg/kg/day exposure.

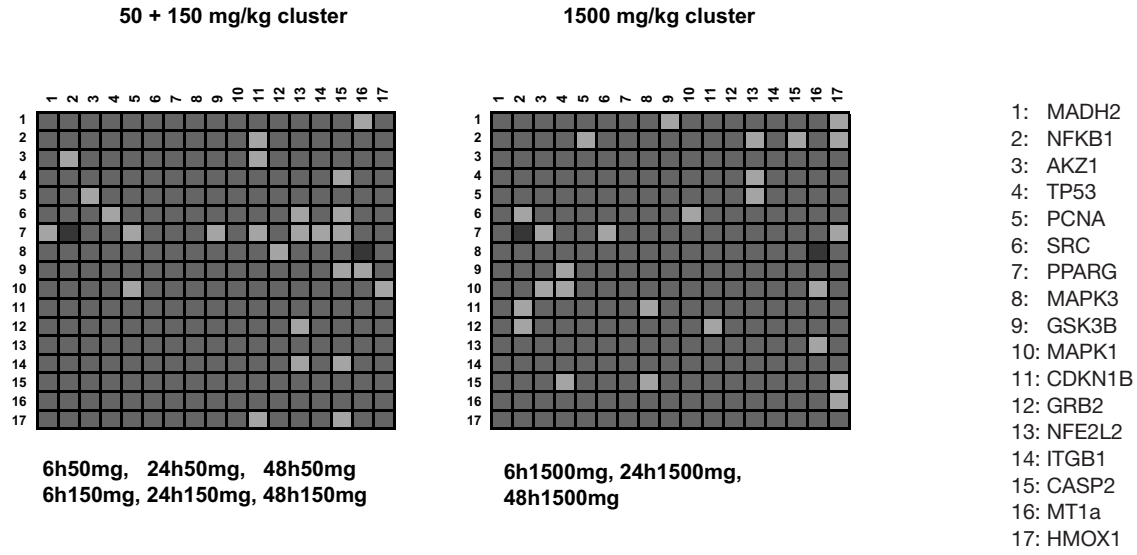


Fig. 2: Two networks resulting from k-means clustering of the estimated networks shown in fig. 1. The left cluster consists of all data for the 50 and 150 mg/kg dose groups and the right cluster consists of all data for the 1500 mg/kg dose group.

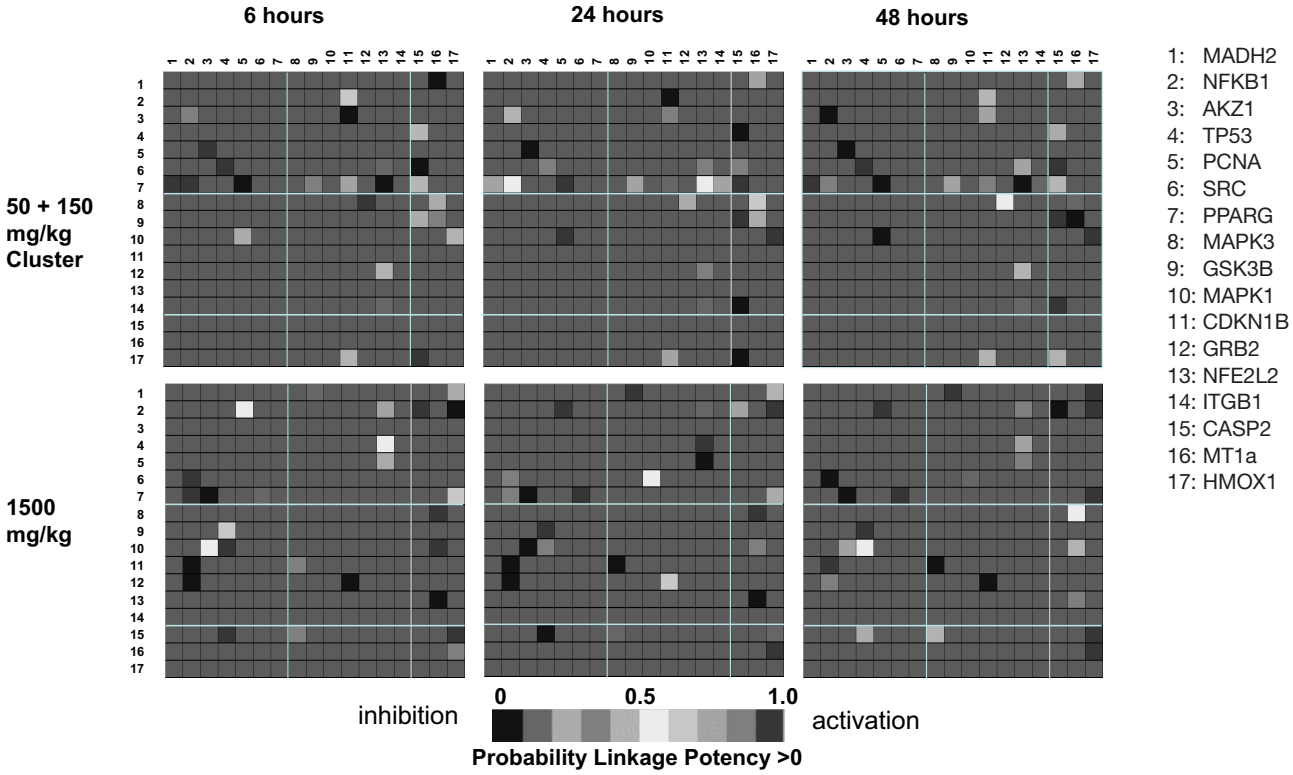


Fig. 3: Significance of the quantified paths in the two centred networks obtained from clustering (fig. 2) after network clustering at each time point. Red indicates significant activation and blue significant inhibition for the path.

networks illustrate significant differences with only two common linkages (shown in red in fig. 2).

To determine the role and the significant level of the linkages between genes in figure 2, a statistical algorithm (Toyoshiba et al., 2004) was applied. To investigate changes over time points, two centred networks were quantified at different time points. The lowest dose (50 mg/kg) and the middle dose (150 mg/kg) data were combined at each time point. Significant levels of each linkage are illustrated in fig. 3. Deeper red and blue represent more significant activator and inhibitor respectively. For example, the linkage from SRC to CASP2 is an inhibition at the earliest time point, changing to activation at later time points. By this quantification analysis, changes in the roles of the linkages in the networks could be illustrated over the time points of the experiment.

Discussion

This study illustrates a new approach with a potential for gaining a mechanistic understanding of chemical toxicity, which is a step further than mere network or clustering analysis. Network type analysis could provide insight into the mechanisms of toxicological responses, but it would be very difficult to interpret if the linkages are increased and complicated. On the other hand, clustering type analysis is useful to compare similarities or differences but does not provide insight into mechanistic changes.

The approach used in this study could provide not only insight into the mechanisms but could also be used to compare networks. In the model experiment with acetaminophen, the high dose group was clustered separately from the low dose group. This might suggest that the network structure at high dose is different from that at the low dose. Also, the quantification analysis could demonstrate how the qualities of the linkages change over time. These results suggest that the network structure may differ between different doses of a substance and that the quality of the linkages may change in the course of exposure to the substance.

To demonstrate another way to use this technique, the transition matrix and the parameter matrix were estimated under low dose acetaminophen, and it was evaluated how well the data from the high dose acetaminophen treatment fitted the model estimated at low dose. Genes that do not fit the model estimated with the low dose data might indicate that the substance has different effects at the higher dose or they might indicate that different linkages exist at the high dose. This information contributes to the understanding of the mechanism of toxicity.

Network structure could be used to characterise chemicals. Studies on one specific chemical of a group could be used to predict the toxicity of other chemicals in the same group. This is an obvious contribution to the 3Rs and may help us to understand the mechanisms of chemical toxicity at lower cost.

References

Beattie, J. H., Owen, H. L., Wallace, S. M., et al. (2005). Metallothionein overexpression and resistance to toxic stress. *Toxicol. Lett. 157(1)*, 69-78.

Bharadwaj, L., Dhami, K., Schneberger, D., et al. (2005). Altered gene expression in human hepatoma HepG2 cells exposed to low-level 2,4-dichlorophenoxyacetic acid and potassium nitrate. *Toxicol. In Vitro 19(5)*, 603-619.

Chen, T., He, H. and Church, G. (1999). Modeling gene expression with differential equations. *Proc. Pac. Symp. On Biocomputing 4*, 29-40.

Friedman, N., Linial, M., Nachman, I. and Pe'er, D. (2000). Using Bayesian networks to analyze expression data. *J. Comput. Biol. 7(3-4)*, 601-620.

Hamadeh, H. K., Knight, B. L., Haugen, A. C. et al. (2002). Methapyrilene toxicity: anchorage of pathologic observations to gene expression alterations. *Toxicol. Pathol. 30(4)*, 470-482.

Harvey, N. L., Butt, A. J. and Kumar, S. (1997). Functional activation of Nedd2/ICH-1 (caspase-2) is an early process in apoptosis. *J. Biol. Chem. 272(20)*, 13134-13139.

Moles, C. G., Mendes, P., Banga and J. R. (2003). Parameter estimation in biochemical pathways: a comparison of global optimization methods. *Genome Res. 13(11)*, 2467-2474.

Pan, H., Zhu, J. and Han, D. (2003). Genetic algorithms applied to multi-class clustering for gene expression data. *Genomics Proteomics Bioinformatics 1(4)*, 279-287.

Panahian, N., Yoshiura, M. and Maines, M. D. (1999). Overexpression of heme oxygenase-1 is neuroprotective in a model of permanent middle cerebral artery occlusion in transgenic mice. *J. Neurochem. 72(3)*, 1187-1203.

Toyoshiba, H., Yamanaka, T., Sone, H. et al. (2004). Gene interaction network suggests dioxin induces a significant linkage between aryl hydrocarbon receptor and retinoic acid receptor beta. *Environ. Health Perspect. 112(12)*, 1217-1224.

Yamanaka, T., Toyoshiba, H., Sone, H. et al. (2004). The TAO-Gen algorithm for identifying gene interaction networks with application to SOS repair in E. coli. *Environ. Health Perspect. 112(16)*, 1614-1621.

Correspondence to

Christopher Portier
Laboratory of Molecular Toxicology
National Institute of Environmental Health Sciences
PO Box 12233 (MD A3-02)
Research Triangle Park
NC, 27709, USA
e-mail: portier@niehs.nih.gov

Theme 7
Applying new science and technology

Plenary 7.0

Using Genomics and Systems Biology to Address Complex Problems: Pancreatic Beta Cell Apoptosis in Diabetes Mellitus

Decio L. Eizirik
Laboratory of Experimental Medicine, Université Libre de Bruxelles, Brussels, Belgium

Summary
In the present text the author briefly discusses the use of the systems biology approach to understand complex human diseases, taking as an example type 1 diabetes mellitus (T1DM). In the course of this discussion, a definition of the systems biology approach is presented and its potential impact on the 3Rs commented.

Keywords: systems biology, microarray analysis, genomics, cytokines

Introduction

It will be discussed here the use of the systems biology approach to understand complex human diseases, taking as an example type 1 diabetes mellitus (T1DM). In the course of this discussion, a definition of the systems biology approach will be presented and its potential impact on the 3Rs commented.

Type 1 diabetes mellitus (T1DM) is one of the most prevalent chronic diseases in children and adolescents, with an incidence ranging from 4-30 cases per 100,000 persons-year in Europe. The global prevalence of T1DM is increasing, and there is no cure for the disease – affected patients require life-long insulin therapy. Unfortunately, exogenous insulin therapy is not a perfect substitute for endogenous insulin production, and the chronic elevation in blood glucose levels may lead to the so called chronic complications of diabetes, including blindness, renal failure, gangrene, myocardial infarction and stroke.

Accumulating evidence indicates that pancreatic beta cells, the only physiological source of insulin production, die by apoptosis in early T1DM. Apoptosis is an active, gene directed process, and recent observations by our group (s. "selected references" below) suggest that beta cell fate following exposure to immune mediators is a complex and highly regulated process, depending on the duration and severity of perturbation of key interacting gene networks. In line with this possibility, genetic susceptibility to T1DM might also depend on gene networks, explaining why it has been so difficult to pinpoint "diabetes susceptibility genes". This departs from the traditional view of phenomena, based on the study of signalling pathways by intuitive inferences based on the study of individual pathway components. Identification of complex and interacting gene/protein patterns poses a formidable challenge, but the sequencing of the human genome, and of the genome of several other species, makes it possible to address it by the use of new high throughput technologies, such as microarray analysis and proteomics. The increasing use of massive parallel analysis of gene/protein expression is emphasising that interpathway cross-talk reflects levels of complexity that cannot be adequately explained by studying individual pathways in isolation. In other words, to fully understand the abnormal cell responses during a pathological stage we need a global multivariate strategy, as proposed by the systems biology approach.

The systems biology approach seeks to devise models based on the comprehensive, qualitative and quantitative analysis of all

constitutive parts of a cell or tissue with the ultimate aim of explaining biological phenomena through the interaction of all its cellular and molecular components. This is based on the analysis of large scale datasets, such as global gene or protein expression. The model is then refined through introduction of perturbations in the system and additional rounds of large scale gene/protein analysis. Systems biology thus turns into an interactive process, in which researchers devise models based on large datasets, make predictions based on the model, and then perform additional large scale experiments to test/validate the prediction and further refine the model.

Against this background, we are utilising microarray analysis and detailed promoter studies to clarify the pattern and regulation of gene expression in primary rat beta cells and in human islets exposed for different time points to the pro-apoptotic cytokines interleukin-1β (IL-1β + interferon-γ (IFN-γ). Nearly 2000 cytokine-induced genes were identified, and the picture emerging from these findings is that beta cells are not passive bystanders of their own destruction. Beta cells respond to cytokine-mediated damage by triggering several genes involved in defense/repair and endoplasmic reticulum stress, by decreasing their most differentiated functions and their capacity for growth and regeneration, and by inducing expression of diverse cytokines and chemokines. Several of these effects of cytokines depend on the activation of the transcription factors NF-κB and STAT-1, and by blocking NF-κB or STAT-1 activation we prevented both cytokine and dsRNA (double stranded RNA) + cytokine-induced rat beta cell death. Subsequent experiments, combining NF-κB blocking and microarray analysis, suggested that NF-κB functions as a "master switch" of IL-1β effects on beta cells, controlling diverse networks of transcription factors and effector genes that contribute to beta cell apoptosis. STAT-1 probably plays a similar role for IFN-g-induced genes. This hypothesis was further investigated by time course and cluster analysis of gene expression in cytokine-treated insulin-producing INS-1 cells, and by "*in silico*" and molecular biology analysis of the promoter regions of genes located in different clusters. Based on the data obtained by our different microarray analysis, we are presently constructing a "Beta Cell Gene Expression Bank", which is already accessible at http://t1dbase.org/cgi-bin/enter_bcgb.cgi. The ultimate goal of this open access resource is to identify and annotate all genes expressed in rat, mouse and human beta cells.

By combining functional studies with microarray analysis, performed with or without targeted perturbations of the system (following the systems biology approach), we hope to eventually understand the complex mechanisms regulating the cytokine-induced beta cell "decision" to undergo apoptosis. This information may point out to new approaches to prevent beta cell death in early T1DM.

These new approaches to understand biological phenomena may have an impact on the 3Rs. Thus:

1. By allowing us to obtain massive information from limited amounts of tissue, it increases the information obtained from a given number of animals, thus decreasing the number or animals required for the experiments. For instance, in the last few years our own laboratory decreased by more than 80% the number of rats and mice utilised in research (these are the only models utilised by our laboratory). A caveat is that in the present phase of exploration and discovery of the genome, there may be an increase in the generation and use of transgenic animals to characterise the function of newly discovered genes.

2. The use of computer models, based on the systems biology approach, may eventually decrease the need of using animals. Computer models of complex and integrated biological systems are as good as the data used to make them. Unfortunately, our present understanding of the interaction between the multiple cellular and molecular components of biological phenomena is limited and fragmentary, making it very difficult to model complex biological responses in a reliable and dynamic way. Thus, additional work on *in vitro* and *in vivo* experimental models is required to accumulate sufficient information for adequate computer modelling. Of note, and in our view, key events detected by computer modelling will, in most cases, need to be confirmed in biological systems.

Conclusion

In conclusion, systems biology is a novel and vibrant field, which is creating new rules as it moves ahead. The challenge it poses is enormous, but it seems that for the first time we have the tools and the adequate experimental approach to tackle biological problems at its real level of complexity.

References

Aitchison, J. D. and Galitski, T. (2003). Inventories to insights. *J. Cell. Biol. 161*, 465-469.

Cardozo, A. K., Berthou, L., Kruhoffer, M. et al. (2003). Gene microarray study corroborates proteomic findings in rodent islet cells. *J. Proteome Res. 2*, 553-555.

Cardozo, A. K., Heimberg, H., Heremenans, Y. et al. (2001). A comprehensive analysis of cytokine-induced and NF-κB dependent genes in primary rat β-cells. *J. Biol. Chem. 276*, 48879-48886.

Cardozo, A. K., Kruhoffer, M., Leeman, R. et al. (2001). Identification of novel cytokine induced genes in pancreatic β-cells by high density oligonucleotide arrays. *Diabetes 50*, 909-920.

Cardozo, A. K., Ortis, F., Feng, Y-M. et al. (2005). Cytokines downregulate the sarcoendoplasmic-reticulum pump Ca^{2+}-ATPase (SERCA)2b and deplete endoplasmic reticulum (ER)-Ca^{2+} leading to induction of ER-stress in pancreatic β-cells. *Diabetes 54*, 452-461.

Eizirik, D. L., Kutlu, B., Rasschaert, J. et al. (2003). Use of microarray analysis to unveil transcription factor and gene networks contributing to β-cell dysfunction and apoptosis. *Ann. N. Y. Acad. Sci. 1005*, 55-74.

Eizirik, D. L. and Mandrup-Poulsen, T. A. (2001). Choice of death – the signal transduction of immune-mediated beta cell apoptosis. *Diabetologia 44*, 2115-2133.

Eldor, R., Yeffet, A., Baum, K. et al. (2006). Conditional and specific NF-kB blockade protects pancreatic beta cells against

diabetogenic inducers. *Proc. Natl. Acad. Sci. USA 103*, 5072-5077.

Gysemans, C., Ladriere, L., Callewaert, H. et al. (2005). Disruption of the IFN-γ signaling pathway at the level of STAT-1 prevents immune destruction of β-cells. *Diabetes 54*, 2396-2403.

Heimberg, H., Heremans, Y., Jobin, C. et al. (2001). Inhibition of cytokine-induced NF-κB activation by adenovirus-mediated expression of a nuclear factor κB super-repressor prevents β-cell apoptosis. *Diabetes 50*, 2219-2224.

Ideker, T., Galitski, T. and Hood, L. (2001). A new approach to decoding life: systems biology. *Annu. Rev. Genomics Human Genet. 2*, 343-372.

Kitano, H. (2002). Systems biology: a brief overview. *Science 295*, 1662-1664.

Kutlu, B., Cardozo, A. K., Darville. M. I. et al. (2003). Discovery of gene networks regulating cytokine-induced dysfunction and apoptosis in insulin-producing INS-1 cells. *Diabetes 52*, 2701-2719.

Kutlu, B., Naamane, N., Berthou, L. and Eizirik, D. L. (2004). New approaches for in silico identification of cytokine-modified β cell gene networks. *Ann. N. Y. Acad. Sci. 1037*, 41-56.

Rasschaert, J., Ladriere, L., Urbain, M. et al. (2005). Toll-like receptor 3 and STAT-1 contribute for dsRNA + IFN-γ-induced apoptosis in pancreatic β-cells. *J. Biol. Chem. 280,* 33984-33991.

Rasschaert, J., Liu, D., Kutlu, B. et al. (2003). Global profiling of double stranded RNA- and INF-γ-induced genes in rat pancreatic beta cells. *Diabetologia 46*, 1641-1657.

Ylipaasto, P., Kutlu, B., Raisilainen, S. et al. (2005). The gene networks regulating coxsackievirus- and cytokine-induced dysfunction and death in primary human beta cells. *Diabetologia*, in press.

Acknowledgements

I thank Dr Nathan Goodman, from the Institute for Systems Biology, Seattle, for helpful discussions. Work by the author was supported by the JDRF Center for Prevention of β-Cell Destruction in Europe under JDRF grant number 4-2002-457, and by grants from the Communaute Française de Belgique – Actions de Recherche Concertees (ARC), Fonds National de la Recherche Scientifique (FNRS) Belgium, European Foundation for the Study of Diabetes (EFSD) and Novo Nordisk Programme in Diabetes Research.

Correspondence to

Decio L. Eizirik, M.D., Ph.D.
Laboratory of Experimental Medicine
Université Libre de Bruxelles
808 Route de Lennik, CP 618
Brussels, B-1070
Belgium
e-mail: deizirik@ulb.ac.be

Session 7.1
Stem cell technology in toxicity testing

A Perspective on Stem Cells as a Tool for *In Vitro* Testing

Aernout Luttun and Catherine M. Verfaillie*
Department of Medicine, Stem Cell Institute, University of Minnesota, Minneapolis, USA

Summary
Stem cells are on their way to becoming a central research tool in this and the next decades. Their differentiation potential promises replacement of lost tissue or provision of functional cells in degenerative diseases. Many of these treatment paradigms have been tested in animals. While stem cells have quickly moved to the frontline of research because they could revolutionise medical practice, they also could be very useful for purposes that do not require animal testing. Here we give an overview on how stem cells can be exploited as in vitro alternative to animals in addressing diverse biological questions.

Keywords: stem cells, differentiation, in vitro testing

Introduction

Like recombinant DNA in the 1980's and transgenic animals in the 1990's, stem cell technology is well on its way to becoming a central research tool in this and the next decades. Stem cells are generally derived from two main sources, embryos or adults, and can be further subcategorised according to their differentiation potential. The latter potential offers great promise for *in vivo* replacement of otherwise irreversible tissue loss seen in many degenerative diseases like myocardial infarction and stroke, or for provision of functional cells in many other disorders where a certain cell type has lost its function, like diabetes and Parkinson's disease. Many of these treatment paradigms have been successfully tested in animal disease models. While stem cells have quickly moved to the frontline of research, mainly because they could revolutionise medical practice, they also could be very useful for other purposes that do not require the use of animals. Here we give an overview on how stem cells in general, and multipotent adult progenitor cells (MAPCs) in particular, can be exploited as *in vitro* alternative to animals in addressing diverse biological questions.

What are stem cells and why could they be a good alternative to animals?

A stem cell is defined by three main criteria: (i) self-renewal, (ii) the ability to differentiate into multiple cells *in vitro* and (iii) the ability to reconstitute a given tissue *in vivo*. Their self-renewal capacity and therefore unlimited availability without the need to sacrifice additional animals for derivation, makes stem cells a particularly useful alternative research tool. The second criterion can be used to divide stem cells into different categories depending on the extent of their differentiation repertoire. On top of the hierarchy is the "totipotent" fertilised egg that not only gives rise to all tissues from the three germ layers within the embryo (ectoderm, mesoderm and endoderm), but also to the supporting trophoblast required for the survival of the developing embryo. Embryonic stem (ES) cells and embryonic germ (EG) cells, isolated from the inner cell mass of the blastocyst or from primordial germ cells of an early embryo, give rise to mesoderm, endoderm, ectoderm and germ cells but not to extra-embryonic tissues, and are therefore termed "pluripotent". Stem cells, isolated from various adult organs, that self-renew and dif-

* Aernout Luttun is a post-doctoral fellow sponsored by the American Heart Association

ferentiate into multiple organ-specific cell types are termed "multipotent stem cells". Committed cells that have limited or no self-renewal ability and differentiate into only one defined cell type are termed "progenitor cells" or "precursor cells". The broader the differentiation potential, the more applications can be considered for a certain type of stem cell.

Pluripotent embryonic or multipotent adult stem cells?

While ES cells are considered pluripotent and therefore offer many potential applications, the differentiation repertoire of adult stem cells is limited and was long believed to be restricted to cell types of the tissue of origin. On the other hand, unlike stem cells from adult sources, the use of ES cells of human origin has met with significant ethical concerns. Over the last 3 years, however, several research groups have challenged the concept of restricted differentiation potential of adult stem cells and have independently identified adult stem cell populations with multi-lineage differentiation that cross lineage restriction boundaries and often include cell types from more than one embryonic germ layer.

We recently described a stem cell with a differentiation capacity that encompasses cell types of all three embryonic germ layers, which we termed "multipotent adult progenitor cell" or MAPC. MAPCs can be isolated from the bone marrow of humans, rodents (Jiang et al., 2002a; Jiang et al., 2002b; Reyes et al., 2002; Reyes et al., 2001), pigs and monkeys (unpublished results). They express the pluripotency marker Oct4 and differentiate into endothelial cells, smooth muscle cells, skeletal myoblasts, neurons, astrocytes, adipocytes, hepatocytes, chondroblasts, and osteoblasts, documented by morphological and functional characterisation (Jiang et al., 2003; Jiang et al., 2002a; Reyes et al., 2002; Schwartz et al., 2002). They also contributed to most of the somatic tissues after injection into the mouse blastocyst (Jiang et al., 2002a). Other groups have subsequently isolated cells from distinct adult tissue sources with differentiation potential that crosses germ layer boundaries (tab. 1). In 2002, Toma et al. isolated "skin derived progenitors" or SKPs. These cells were derived from the dermis of foetal and adult rodent skin, from human scalp and recently also from neonatal human foreskin (Toma et al., 2001; Toma et al., 2005). SKPs differentiate into neurons, glia, as well as smooth muscle cells and mesenchymal lineage cells (osteoblasts and adipocytes). Two years later, MIAMI (marrow-isolated adult multilineage inducible) cells were described, derived from the bone marrow and able to differentiate into cells of all three germ layers, including osteoblasts, chondrocytes, adipocytes, neurons, and pancreatic islet-like cells (D'Ippolito et al., 2004). Kogler et al. also found a multipotent stem cell population in human cord blood, called "unrestricted somatic stem cell" or USSC, contributing to cell types from the three germ layers *in vitro* and *in*

Tab. 1: Overview of multipotent adult stem/progenitor cells

Name	origin	year	OCT4 expression	FACS phenotype	Differentiation		
					mesoderm	ectoderm	endoderm
SKP	rodent skin human scalp human foreskin	2002 2005	ND	ND	yes	yes	ND
MAPC	rodent, human, pig, monkey BM rodent brain rodent skeletal muscle	2002	yes	*Mouse:* $CD13^+SSEA1^+Flk1^{low}CD90^{low}Sca1^{low}$ $CD19^-CD3^-CD34^-CD45^-CD44^-CD117^-Gr1^-$ $Mac1^-$MHCclass I^-MHCclass II^- *Human:* $CD13^+Flt1^{low}KDR^{low}CD133^{low}$β–micro-$globulin^{low}CD31^-CD34^-CD45^-CD44^-CD36^-$ $CD106^-$VE-$Cadherin^-CD62E^-CD62L^-CD62P^-$ $Tek^-Tie^-CD117^-$glyA$^-$HLA-ABC$^-$HLA-DR$^-$	yes	yes	yes
MIAMI	human BM	2004	yes	$CD29^+CD49^+CD63^+CD81^+CD90^+CD122^+$ $CD164^+CD36^-CD54^-CD56^-CD117^-CD45^-$ $CD34^-$HLA-DR$^-$	yes	yes	yes
USSC	human cord blood	2004	ND	$CD13^+CD29^+CD44^+CD49e^+CD90^+CD105^+$ $CK8^+CK18^+CD10^{low}KDR^{low}$ HLA-$ABC^{low}CD14^-CD33^-CD34^-CD45^-$ $CD49b^-CD49c^-CD49d^-CD49f^-CD50^-CD62E^-$ $CD62L^-CD62P^-CD106^-CD117^-$glyA$^-$HLA-DR$^-$	yes	yes	yes
BMSC	human BM	2005	no	$KDR^-CD90^-CD105^-CD117^-CD29^-CD44^-$ $CD73^-CD45^-CD49f^-CD50^-CD34^-CD133^-$ HLA-ABC^-HLA-DR$^-$	yes	yes	yes

SKP: skin-derived precursors; MAPC: multipotent adult progenitor cells; MIAMI: marrow-isolated adult multilineage inducible; USSC: unrestricted somatic stem cells; BMSC: BM-derived multipotent stem cells; BM: bone marrow; CK: cytokeratin; glyA: glycophorin A; ND: not documented

vivo (Kogler et al., 2004). Finally, last year, Losordo's group in Boston isolated multipotent cells from human bone marrow with a differentiation potential similar to MAPCs, but with different surface expression of several markers and lacking expression of the pluripotency marker Oct4 (Yoon et al., 2005). The broad differentiation potential of all these adult stem cell types along with their availability without ethical restrictions offers numerous possible applications *in vitro* and *in vivo*. Unlike ES or EG cells, none of these cell types have so far been documented to contribute to the germline.

How can (stem) cells be used as alternatives to animals?

Many *in vitro* techniques have been developed using different resources, including microorganisms, plants, invertebrates (i.e. *Hydra attenuata*), and embryos, tissues or cells from vertebrate animals and humans (Pearson, 1986). The best-known application of cells as an alternative to animal usage is the toxicity testing of drugs and chemicals. While numerous somatic cell lines (i.e. human keratinocytes for phototoxicity, lung cells for inhalation toxicity, etc.) have proven useful to test the toxicity or sensitivity to certain compounds, ES cells from animals are of particular interest in developmental toxicity testing. The embryonic stem cell test (EST) takes advantage of the potential of murine embryonic stem (ES) cells to differentiate in culture to test embryo toxicity *in vitro* and to screen for teratogenicity and growth retardation. The EST represents an *in vitro* system for the classification of compounds according to their teratogenic potential based on the morphological analysis of beating cardiomyocytes in embryoid body (EB) outgrowths compared to cytotoxic effects on undifferentiated murine ES cells and differentiated 3T3 fibroblasts. The test has recently been refined by defining more objective endpoints of differentiation other than the microscopic evaluation of "beating areas", such as assessment by fluorescence activated cell sorting (FACS) and quantitative reverse transcription polymerase chain reaction (qRT-PCR), which evaluate expression of genes at the translational and transcriptional level, respectively (Seiler et al., 2004). In addition, in order to address the potential concern of interspecies variations in developmental toxicity, some considerations have been made to adapt the mouse ES cell based tests to human ES cells (Pellizzer et al., 2005).

While embryonic stem cells have been very useful for *in vitro* developmental toxicity testing, *in vitro* adult stem cell research in general and testing with haematopoietic stem cells in particular, has been of tremendous importance for the unravelling of proliferation and differentiation mechanisms and for the discovery of several haematopoietic cytokines, which are now used worldwide for the treatment of (cancer) patients undergoing chemotherapy and radiation therapy. Two examples are the recombinant proteins erythropoietin and filgrastim (Neupogen).

Another major domain of cell-based *in vitro* testing is the biotransformation of drug compounds. Primary hepatocyte cell lines have frequently been used for that purpose. Gastric epithelial cells (i.e. Caco-2 cells) and co-cultures of brain microvessel endothelial cells with glial cells have been used for bioavailability and drug transport studies. Neurons derived from the embryonic rat spinal cord and dorsal root ganglia have been used as a supplement to *in vivo* testing of analgesic drugs. While the creation of knock-out animals has advanced the search for disease mechanisms tremendously, *in vitro* alternatives have now been developed, such as antisense nucleic acids or oligonucleotides and RNA interference to knock-down gene expression in specific cell targets *in vitro.* For a more extensive database of possible applications of cells as an alternative, the reader is referred to the following website: http://gripsdb.dimdi.de.

One major disadvantage of working with primary cell cultures is their limited proliferation ability, which necessitates repeated isolation and sacrifice of additional animals for their derivation. One possible solution to this problem is to immortalise the cells, which may however change their characteristics and function significantly. As mentioned above, stem cells are able to self-renew indefinitely, and therefore can be expanded without limitations. For many of the cell types mentioned above, adequate differentiation protocols have been established starting from stem cells, both embryonic and adult stem cells. Thus, stem cells could serve as a continuous source of such specialised cell types. When establishing a differentiation protocol, it is very important to document functionality of the resulting cell type. Functional cell differentiation has been elaborately documented for MAPC-derived cells such as endothelial cells (Reyes et al., 2002), hepatocytes (Schwartz et al., 2002), neurons (Jiang et al., 2003) and smooth muscle cells (unpublished results). Although *in vivo* transplantation into animal models is the most robust method to ultimately prove functionality, there are adequate *in vitro* tests available to document function. Endothelial function, for instance, was shown by uptake of acetylated LDL, by release of von Willebrand factor (vWF) upon histamine stimulation and by the formation of vascular tubes in a two-dimensional matrix (matrigel) system. Liver cell function was demonstrated by production of albumin and urea, by cytochrome P450 induction, glycogen storage and uptake of LDL. Neurons differentiated from MAPCs were shown to be functional by their electrophysiological properties similar to midbrain-derived neurons. Finally, (vascular) smooth muscle cell function is tested by responsiveness to smooth muscle cell agonists and by production of extracellular matrix components (fibrillar collagen and elastin). For other cell types, such as cardiomyocytes, differentiation has been suggested by the presence of MAPC-derived cells in the heart following blastocyst injection, but the conditions for *in vitro* (functional) differentiation require further optimisation. Functional cardiomyocyte differentiation has been documented in ES cells (human and rodent), where function is assessed by electrophysiology (Boheler et al., 2004; Kogler et al., 2004; Lakshmipathy et al., 2004). A limitation to differentiation from ES cells is that in most cases differentiation requires the formation of three-dimensional structures (embryoid bodies or EBs), which often produce a mix of differentiated cell types instead of a pure culture of one differentiated cell type. Genetic manipulation protocols, using cell-specific promoters that drive a fluorescent colour, have been developed to generate purified cell populations by FACS sorting.

Examples of *in vitro* applications using MAPCs

In the following paragraph, we demonstrate how MAPCs or cells derived from MAPCs can be used for several *in vitro* applications. We also show how two different disciplines of science, stem cell biology and bioengineering, can join forces to find answers to biological questions.

Example 1: Creating artificial vascular grafts for bypass surgery

Myocardial infarction is one of the leading causes of morbidity and mortality worldwide. Its underlying cause is atherosclerotic degeneration of arteries that normally supply the heart muscle with oxygenated blood. As a result of inadequate arterial blood supply, part of the heart muscle dies and is replaced by dysfunctional scar tissue. The area around the dead tissue, however, can be potentially prevented from dying by restoring blood supply to the heart. Stem cells have now been tested for their *in vivo* ability to induce the formation of new blood vessels (arteries) and to replace the dead cardiac muscle by new cardiomyocytes, both in animal models and in patients. Although the initial results show promise, it is unlikely that stem cells alone will be sufficient to treat every single patient. Rather, stem cell transplantation could be a useful addition to and be combined with already existing approaches such as bypass surgery. The latter technique restores blood flow to the oxygen-deprived heart muscle by replacing one or more segments of the atherosclerotic arteries by vascular grafts, usually taken from the veins of the patient. Stem cells could be useful to create artificial vessel grafts. The latter would require differentiation of the stem cells into endothelial cells and vascular smooth muscle cells and their correct positioning following differentiation (i.e. endothelial cells lining the inside of the graft and smooth muscle cells surrounding the endothelial layer). Importantly, since these grafts need to be implanted into arteries, it would be appropriate to generate endothelial cells that have characteristics and functions of arterial endothelial cells. Indeed, it has recently been documented that endothelial cells from arteries and those from veins express different sets of genes, some of which may translate into different functions (Chi et al., 2003; Torres-Vazquez et al., 2003). Therefore, one of the current challenges in vascular repair and regeneration is to develop protocols that generate arterial endothelial cells. We have recently shown the ability of human MAPCs to be induced towards an arterial endothelial phenotype (unpublished results in collaboration with F. Prosper, University of Navarra, Spain). We also successfully generated artificial grafts in which human MAPC derived ECs lined the inside of a vascular graft made up of rat aortic smooth muscle cells (unpublished results in collaboration with B. Tranquillo, University of Minnesota, Minneapolis). Finally, appropriate communication between smooth muscle cells and endothelial cells may be required to generate vascular grafts as they may co-influence their proper differentiation. Indeed, it has been shown that co-transplantation of endothelial and smooth muscle cells results in the formation of much more stable, functional and durable vessels *in vivo* than transplantation of endothelial cells alone (Koike et al., 2004). We recently developed a three-dimensional matrix system in which both cell types can be co-cultured to study their interaction (unpublished results).

Example 2: Three-dimensional tissue engineering

For a system to be as adequate as animals to study certain biological phenomena, it should mimic the *in vivo* microenvironment as closely as possible. One of the goals of bioengineering is to reconstruct tissue according to its native architecture. All tissues in a living higher organism are dependent on blood flow for their proper oxygenation and nutrition. Moreover, the presence of blood vessels and vascular cell-derived molecules has been shown to be essential for tissue development and differentiation, as shown for the pancreas and the liver (Cleaver and Melton, 2003; LeCouter et al., 2003). Creating three-dimensional artificial liver constructs *in vitro* could be a valuable alternative to animals for drug toxicity testing and for biotransformation studies. One of the structural elements of the liver is the sinusoid, composed of a fenestrated vascular structure decorated with hepatocytes. In the adult liver, sinusoids are aligned radially, running from the lobule periphery to the central vein. This sinusoidal organisation is important for the function of the liver. Recently, the group of D. Odde (University of Minnesota, Minneapolis) developed a technique called "laser-guided direct writing" (LGDW). Using this method, they were able to pattern multiple cell types with micrometer precision onto biological matrices (Nahmias et al., 2004; Odde and Renn, 2000). In a first approach, they could lay down MAPCs into a matrix in a three-dimensional pattern without significant negative effect on cell viability (Nahmias et al., 2004). Subsequently, in an effort to recreate the liver tissue architecture *in vitro*, they patterned human umbilical cord endothelial cells (HUVECS) in several linear formations. Like for MAPCs, this did not compromise the viability of the endothelial cells but instead stimulated tube formation. Co-culture of these vascular tubes with hepatocytes resulted in an aggregated tubular structure, similar in organisation to a hepatic sinusoid (Nahmias et al., 2005). As the development of liver sinusoids takes several weeks *in vivo,* these endothelial-hepatic aggregates formed after a few days in culture are likely not fully developed sinusoids. Nevertheless, these results suggest that LGDW can be used to create a vascular backbone for *ex vivo* liver morphogenesis. Alternatively, LGDW could be used to lay down beads coated with different growth factors onto stem cells in order to induce their differentiation in a specific spatial pattern.

Conclusion

Many academic institutions have become aware of the excessive use of animals in their research programmes and have therefore started to implement policies to reduce the number of animals, to refine methodologies to diminish pain and suffering of animals and to encourage the development of alternative research tools that can answer biological questions equally well. Stem cell technology is one of these alternatives. The unlimited availability of stem cells from different species and our increasing knowledge on how to generate functional cell types from them

opens many possibilities for the design of stem cell-based *in vitro* test protocols, some of which are documented in this short overview.

References

Boheler, K. R., Crider, D. G., Tarasova, Y. and Maltsev, V. A. (2004). Cardiomyocytes derived from embryonic stem cells. *Methods Mol. Med. 108*, 417-436.

Chi, J. T., Chang, H. Y., Haraldsen, G. et al. (2003). Endothelial cell diversity revealed by global expression profiling. *Proc. Natl. Acad. Sci. USA 100*, 10623-10628.

Cleaver, O. and Melton, D. A. (2003). Endothelial signaling during development. *Nat. Med. 9*, 661-668.

D'Ippolito, G., Diabira, S., Howard, G. A. et al. (2004). Marrow-isolated adult multilineage inducible (MIAMI) cells, a unique population of postnatal young and old human cells with extensive expansion and differentiation potential. *J. Cell. Sci. 117*, 2971-2981.

Jiang, Y., Henderson, D., Blackstad, M. et al. (2003). Neuroectodermal differentiation from mouse multipotent adult progenitor cells. *Proc. Natl. Acad. Sci. USA 100 Suppl 1*, 11854-11860.

Jiang, Y., Jahagirdar, B. N., Reinhardt, R. L. et al. (2002a). Pluripotency of mesenchymal stem cells derived from adult marrow. *Nature 418*, 41-49.

Jiang, Y., Vaessen, B., Lenvik, T. et al. (2002b). Multipotent progenitor cells can be isolated from postnatal murine bone marrow, muscle, and brain. *Exp. Hematol. 30*, 896-904.

Kogler, G., Sensken, S., Airey, J. A. et al. (2004). A new human somatic stem cell from placental cord blood with intrinsic pluripotent differentiation potential. *J. Exp. Med. 200*, 123-135.

Koike, N., Fukumura, D., Gralla, O. et al. (2004). Tissue engineering: creation of long-lasting blood vessels. *Nature 428*, 138-139.

Lakshmipathy, U., Pelacho, B., Sudo, K. et al. (2004). Efficient transfection of embryonic and adult stem cells. *Stem Cells 22*, 531-543.

LeCouter, J., Moritz, D. R., Li, B. et al. (2003). Angiogenesis-independent endothelial protection of liver: role of VEGFR-1. *Science 299*, 890-893.

Nahmias, Y., Schwartz, R. E., Verfaillie, C. M. and Odde, D. J. (2005). Laser-guided direct writing for three-dimensional tissue engineering. Biotechnol. Bioeng. (publication ahead of print).

Nahmias, Y. K., Gao, B. Z. and Odde, D. J. (2004). Dimensionless parameters for the design of optical traps and laser guidance systems. *Appl. Opt. 43*, 3999-4006.

Odde, D. J. and Renn, M. J. (2000). Laser-guided direct writing of living cells. *Biotechnol. Bioeng. 67*, 312-318.

Pearson, R. M. (1986). In-vitro techniques: can they replace animal testing? *Hum. Reprod. 1*, 559-560.

Pellizzer, C., Bremer, S. and Hartung, T. (2005). Developmental toxicity testing from animal towards embryonic stem cells. *ALTEX 22*, 47-57.

Reyes, M., Dudek, A., Jahagirdar, B. et al. (2002). Origin of endothelial progenitors in human postnatal bone marrow. *J. Clin. Invest. 109*, 337-346.

Reyes, M., Lund, T., Lenvik, T. et al. (2001). Purification and ex vivo expansion of postnatal human marrow mesodermal progenitor cells. *Blood 98*, 2615-2625.

Schwartz, R. E., Reyes, M., Koodie, L. et al. (2002). Multipotent adult progenitor cells from bone marrow differentiate into functional hepatocyte-like cells. *J. Clin. Invest. 109*, 1291-1302.

Seiler, A., Visan, A., Buesen, R. et al. (2004). Improvement of an in vitro stem cell assay for developmental toxicity: the use of molecular endpoints in the embryonic stem cell test. *Reprod. Toxicol. 18*, 231-240.

Toma, J. G., Akhavan, M., Fernandes, K. J. et al. (2001). Isolation of multipotent adult stem cells from the dermis of mammalian skin. *Nat. Cell. Biol. 3*, 778-784.

Toma, J. G., McKenzie, I. A., Bagli, D. and Miller, F. D. (2005). Isolation and characterization of multipotent skin-derived precursors from human skin. *Stem Cells 23*, 727-737.

Torres-Vazquez, J., Kamei, M. and Weinstein, B. M. (2003). Molecular distinction between arteries and veins. *Cell Tissue Res. 314*, 43-59.

Yoon, Y. S., Wecker, A., Heyd, L. et al. (2005). Clonally expanded novel multipotent stem cells from human bone marrow regenerate myocardium after myocardial infarction. *J. Clin. Invest. 115*, 326-338.

Correspondence to

Catherine M. Verfaillie
Department of Medicine, Stem Cell Institute
University of Minnesota
2001 SE 6th Street
MN55455 Minneapolis, USA
e-mail: aernout_luttun@yahoo.com

Current Status of the Embryonic Stem Cell Test: The Use of Recent Advances in the Field of Stem Cell Technology and Gene Expression Analysis

Andrea Seiler, Roland Buesen, Katrin Hayess, Katharina Schlechter, Anke Visan, Elke Genschow, Birgitta Slawik and Horst Spielmann

Federal Institute for Risk Assessment (BfR), National Center for Documentation and Evaluation of Alternative Methods to Animal Experiments (ZEBET), Berlin, Germany

Summary

All guidelines that are currently used for regulatory developmental toxicity testing of chemicals and drugs are based on animal experiments. The most promising alternative is based on embryonic stem cells of the mouse (mESC). Their ability to differentiate into numerous cell types have made ES cells a popular system to study gene function and developmental processes during differentiation in vitro. The embryonic stem cell test (EST) makes use of this capacity to detect developmental toxicants during differentiation of stem cells into cardiomyocytes.

In the present study our investigations were aimed at the further development of the validated EST protocol. We present improvements that focus primarily on (i) the quantitative assessment of drug effects at the cellular level, using a novel approach in which the expression of tissue-specific marker proteins under the influence of the test chemical is quantified by intracellular flow cytometry in ES cells, (ii) the development of protocols for ES cell differentiation into various cell types other than cardiomyocytes, e.g. neural cells, and (iii) the standardisation and optimisation of ES cell culture and differentiation conditions in chemically defined serum-free medium. An important strength of the molecular approach in combination with serum-free culture conditions is that in this way the ability of the test to monitor the cellular response to toxins could be expanded to proteins of many signal transduction pathways in a highly standardised form. Furthermore, these improvements now allow testing of substances known to interact with serum proteins.

Keywords: in vitro, embryotoxicity, embryonic stem cell test, differentiation, cytotoxicity, molecular endpoints, intracellular flow cytometry, developmental neurotoxicity testing, serum replacement

Introduction

Assessing the toxicity of chemicals for development and the reproductive cycle according to standardised OECD Test Guidelines requires extensive screening and multi-generation studies (OECD, 1983, 1996, 2001a, 2001b, 2001c). For chemicals used as drugs, "segment studies" covering three important phases of pre- and postnatal development including fertility have to be conducted (ICH, 1993). These *in vivo* test methods are time-consuming, expensive and have to be carried out on large numbers of laboratory animals. For example, a developmental toxicity study (OECD TG 414) requires 150 animals and a two-generation study (OECD TG 416) more than 3,000 animals, including pups.

According to the currently ongoing implementation of the new chemicals policy of the European Union (REACH, **R**egistration, **E**valuation, **A**uthorisation of **Ch**emicals) the toxicity of approximately 30,000 existing chemicals produced or marketed at more than one tonne per year has to be evaluated in the European Union within a period of 12-15 years. The procedures for assessing reproductive toxicity will have the strongest impact on the total number of animals used for testing under REACH (Höfer et al., 2004; see also Anon., 2005). Furthermore it has been estimated that the testing costs for the developmental toxicity study and for the two generation study will represent up to 57% of the total testing costs of REACH (van der Jagt et al., 2004; see also Eskes and Zuang, 2005). In addition, the 7th Amendment of the Cosmetic Directive demands fixed deadlines for phasing out animal experiments for safety toxicity testing including reproductive toxicity (Anon., 2003). Thus, predictive screens for the evaluation of reproductive toxicity need to be made available with the ultimate goal of reducing animal use and testing more chemicals than can be accommodated by conventional whole animal testing.

To obtain information on the toxic effects of chemicals and drugs on specific elements of the reproductive cycle, reproductive toxicity testing is either performed *in vivo* using pregnant animals or *in vitro* on cultured embryos or embryonic cells and tissues from pregnant animals. Both for *in vivo* and for *in vitro* testing pregnant animals have to be sacrificed to obtain embryonic cells, tissues or organs. Taking advantage of the potential of ESC to differentiate in culture into a variety of cell types (reviewed in Kirschstein and Skirboll, 2001; Smith, 2001) an *in vitro* developmental toxicity test with two permanent cell lines from the mouse has been proposed, i.e. the *Embryonic Stem Cell Test* (EST) (Spielmann et al., 1997). Using the ability of mESC

to differentiate spontaneously into contracting cardiomyocytes, a reliable assay system has been designed for the assessment of embryotoxic/teratogenic properties of chemicals and drugs *in vitro*.

The validated EST

Murine ESC of the permanent cell line D3 (Doetschmann et al., 1985) are maintained in an undifferentiated state in culture in the presence of the leukemia inhibitory factor (mLIF; Williams et al., 1988). Differentiation of mESC is induced by the withdrawal of mLIF. Using the "hanging drop" culture technique described by Rudnicki and McBurney (1987) mESC form multicellular aggregates called embryoid bodies (EBs). Within these aggregates, complex interactions between heterologous cell types result in the induction of differentiation of stem cells to derivatives of all three embryonic germ layers (Martin et al., 1977), including contracting myocardial cells (Doetschmann et al., 1985; Rudnicky and McBurney, 1987; Maltsev et al., 1994; Hescheler et al., 1997; Guan et al., 1999; Boheler et al., 2002; Sachinidis et al., 2003).

The EST benefits from the fact that differentiation into contracting cardiomyocytes can be detected easily by microscopic inspection of EB outgrowths at day 10 of differentiation. In addition to the differentiation analysis, cytotoxic effects of the test substance on mESC and 3T3 fibroblasts are assessed (fig. 1). By using mESC and differentiated (adult) fibroblasts, the assay takes embryonic as well as maternal toxicity into account. To assess the concentration of a substance that inhibits differentiation (ID_{50}) and proliferation (IC_{50}D3 and IC_{50}3T3) by 50% compared to the untreated control, concentration-response profiles are generated. To predict the toxic potential of a test compound *in vivo* from *in vitro* data, a biostatistical prediction model (PM) was developed to assign the test compounds to three classes of embryotoxicity: *non-embryotoxic, weakly embryotoxic and strongly embryotoxic* (Genschow et al., 2002 and 2004).

In an international ECVAM validation study on three *in vitro* embryotoxicity tests, the EST was scientifically validated using a set of 20 reference compounds characterised by high-quality *in vivo* embryotoxicity data assessed in laboratory animals and humans. The EST predicted the embryotoxic potential of the 20 reference compounds with an accuracy of 78%. Remarkably, a predictivity of 100% was obtained for strongly embryotoxic chemicals (Genschow et al., 2002 and 2004). According to the ECVAM Scientific Advisory Committee (ESAC), the three *in vitro* methods for embryotoxicity testing (EST, micromass test; and postimplantation rat whole-embryo culture test) are scientifically validated and ready for consideration for regulatory acceptance and application (Balls and Hellsten, 2002).

Current improvements of the EST

Since the EST has been demonstrated to be a reliable alternative method for embryotoxicity testing, several attempts have been made by our group and others to improve the EST protocol. Our recent improvements were predominantly focused on the following objectives: (i) the quantitative assessment of drug effects at the cellular level to allow a more objective measurement, (ii) the establishment of additional tissue-specific endpoints to increase the precision of the assay and (iii) the replacement of foetal calf serum (FCS) in the EST to further standardise the test system.

The successful establishment of new molecular endpoints in the EST, including the expansion to further cell-type specific endpoints like neural cells and the successful implementation of serum-free culture conditions in the EST were presented in an oral presentation at the 5th World Congress on Alternatives & Animal Use in the Life Sciences in Berlin, Germany (September 21st-25st, 2005).

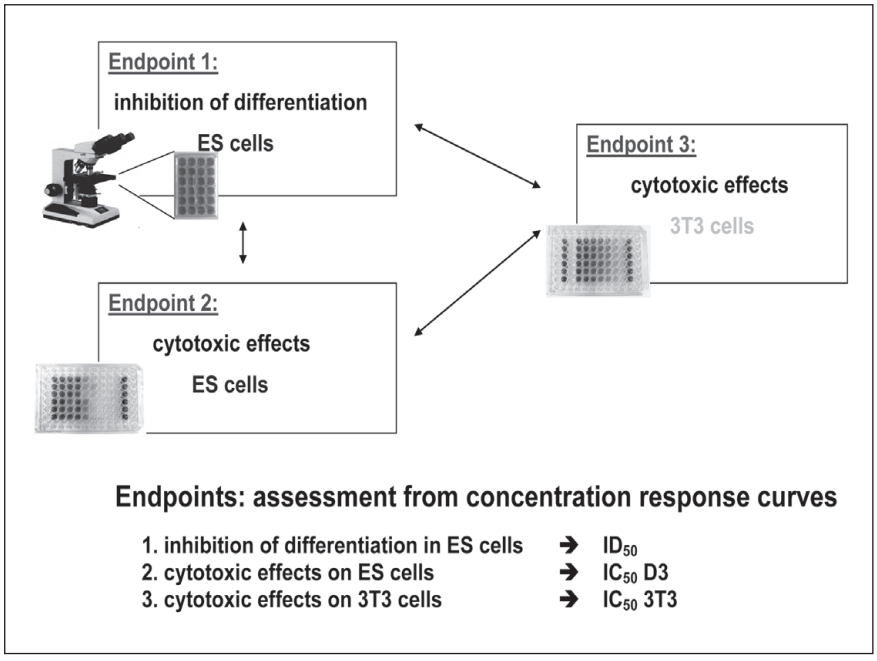


Fig. 1: The embryonic stem cell test (EST). Schematic overview illustrating the principle approach and the endpoints applied in the test to assess the embryotoxic potential of test compounds using two permanent mouse cell lines: 3T3 fibroblasts and mESC (cell line D3).

Methods

Murine ESC were differentiated into beating cardiomyocytes using the hanging drop method (Rudnicki and McBurney, 1987). Test substances were applied throughout differentiation. Differentiation was determined by microscopic analysis of beating cardiomyocytes at day 10 of development (validated endpoint) and by quantitative gene expression analysis of selected tissue-specific marker genes using flow cytometry (day 7 of differentiation, new molecular endpoint). Concentration-response curves were generated to determine the concentration of a substance which inhibits the development of contracting cardiac muscle cells by 50% (ID_{50}). Efficient differentiation into contracting cardiomyocytes from mESC (cell line D3), chemical treatment and the quantitative intracellular flow cytometry approach were performed as described previously (Seiler et al., 2004, Seiler et al., 2006, in press).

For neural cell differentiation adherent monocultures in a defined medium according to Ying and Smith (2003), were generated with some modifications (K. Hayess, A. Visan and A. Seiler unpublished data).

To investigate cardiac ESC differentiation under serum-free conditions, mESC were adapted to serum-free medium in steps by raising the concentration of the serum replacement (Gibco, Karlsruhe, Germany) supplemented with specific factors in the medium (Advanced-DMEM; Gibco, Karlsruhe, Germany) while simultaneously reducing the foetal calf serum (FCS) concentration (Schlechter et al., manuscript in preparation). The adapted mESC were differentiated into beating cardiomyocytes according to the validated EST protocol (Spielmann et al., 1997; Seiler et al., 2006, in press).

Results

Improvements of the Embryonic Stem Cell Test (EST)

Implementing molecular endpoints of differentiation

In order to identify more objective endpoints other than the microscopic evaluation of "beating areas" and to follow cardiac differentiation at the cellular level, we improved and expanded the EST protocol by establishing molecular endpoints of differentiation. The quantitative expression of sarcomeric myosin heavy chain (MHC) and α-actinin genes under the influence of selected reference compounds was studied employing intracellular flow cytometry (Seiler et al., 2004). These results indicated that structural proteins of the sarcomere apparatus are promising candidates to predict developmental toxicity *in vivo* from *in vitro* data (Seiler et al., 2004).

To prove the general applicability of the new molecular endpoint in comparison to the validated microscopic evaluation, we are currently performing a prevalidation study using a set of ten chemicals with different embryotoxic potentials. For each chemical the new molecular endpoint and the validated endpoint – the microscopic analysis of beating cardiomyocytes – were assessed from concentration response curves. Results for two strongly embryotoxic, two weakly embryotoxic and one non-embryotoxic compound are presented in figure 2. Almost identical ID_{50} values were obtained with both methods. These data indicate that the molecular approach is as predictive as the validated endpoint and that gene expression analysis using intracellular flow cytometry can serve as a new toxicological endpoint in the EST (Seiler et al., 2004; Buesen et. al., manuscript in preparation).

Differentiation of mESC into neural cells

The current experimental design of the EST involves differentiation of mESC into contracting cardiomyocytes. However, potentially embryotoxic drugs may target tissues other than the myocardium. This consideration prompted us to expand the EST to other major target tissues. To create an efficient *in vitro* approach for developmental neurotoxicity, we established a protocol for the differentiation of mESC into neural cells designed with special regard to the testing of chemicals and drugs.

In the past, for neural cell differentiation from pluripotent embryonic stem cells we followed the protocol published by Okabe et al. (1996) (lineage selection) with minor modifications or stimulated the induction of neural cells by retinoic acid treatment (Bain et al., 1996). Although the protocols were successfully established in our lab, they are not ideal for use in *in vitro*

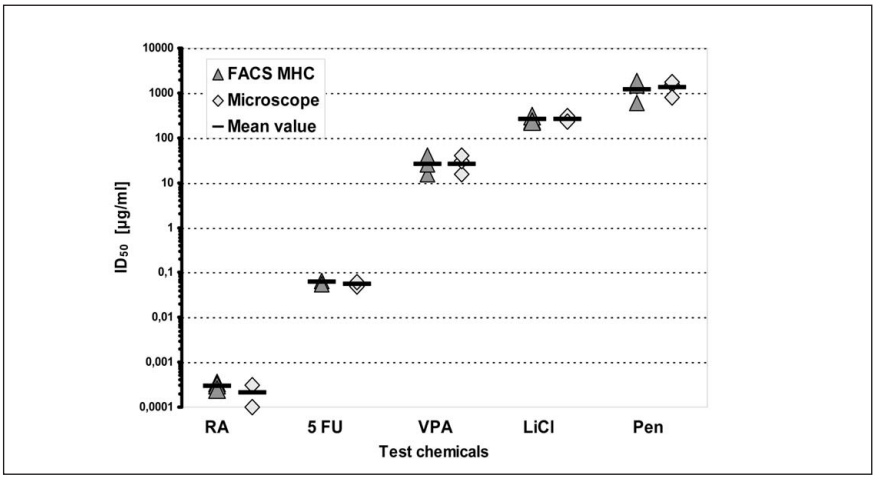


Fig. 2: Comparison of ID_{50} values. ID_{50} values (n=3) were assessed for each chemical from concentration response curves using the validated differentiation endpoint (microscopic evaluation of beating areas at day 10) or the molecular endpoint, quantification of the cardiac-specific marker protein myosin heavy chain (MHC) by intracellular flow cytometry on day 7. Abbreviations: RA, all-trans-retinoic acid; 5 FU, 5-Fluorouracil; VPA, valproic acid; LiCl, lithium chloride; Pen, penicillin G.

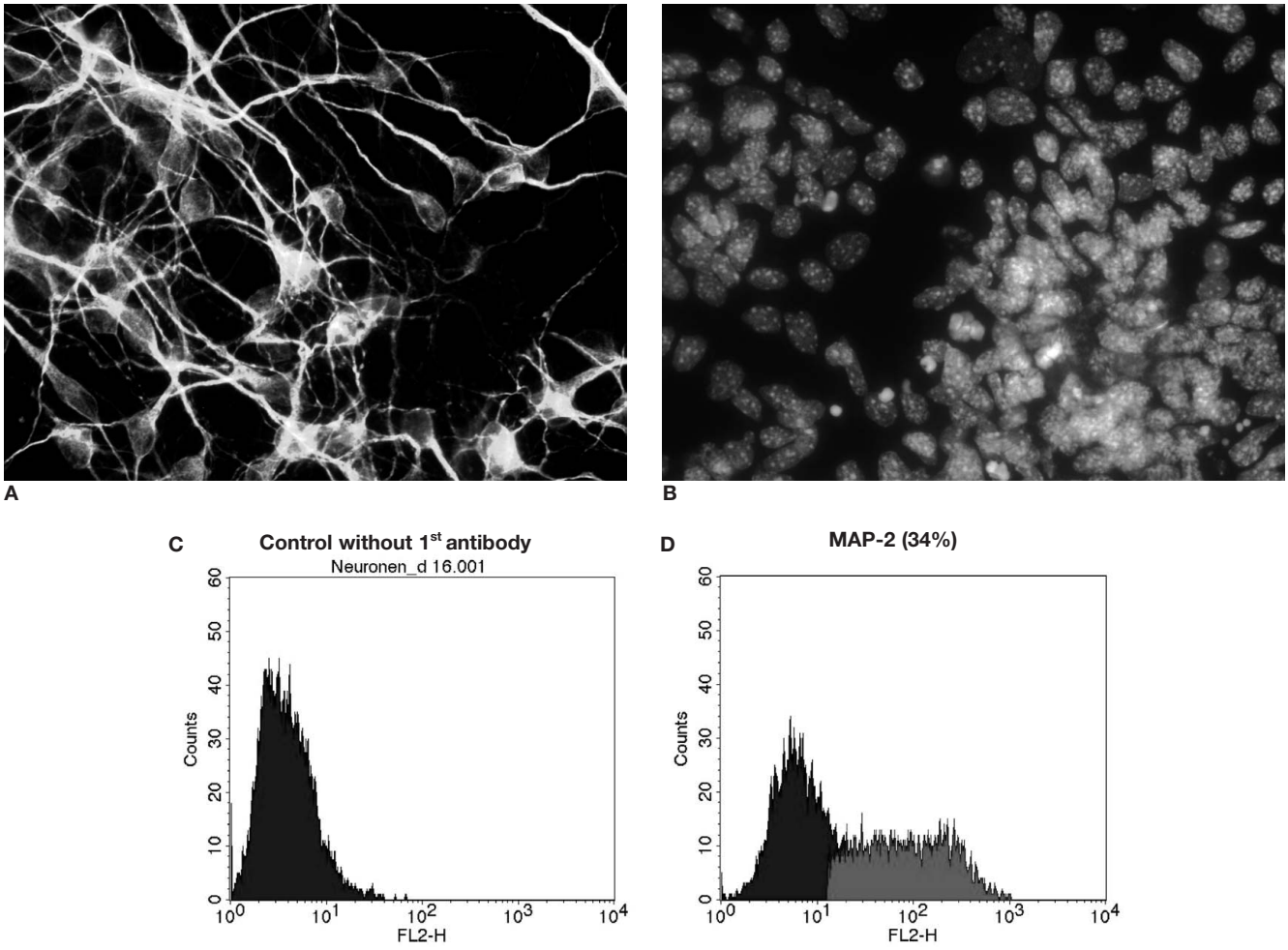


Fig. 3: Differentiation of neuronal cells from mESC using adherent monolayer cultures and quantification of neural marker protein expression by intracellular flow cytometry. Representative immunofluorescence staining of neural marker protein with anti-MAP-2, stained at day 16 of differentiation (A) with corresponding nuclear DAPI (4',6-diamino-2-phenylindole) staining. (B) Representative flow cytometry analysis of anti-MAP-2 stained cells at day 16 of differentiation (murine ES cell line D3). The x-axis corresponds to the fluorescence intensity and the y-axis to the number of cells per channel (events) (D). Control lacking primary antibody (C).

developmental neurotoxicity assays. The differentiation times are quite long and the growth factors are quite expensive. In the case of the second protocol mentioned above, the inducer substance, retinoic acid, is itself a strongly embryotoxic compound. To find an alternative, we studied the protocol published by the group of Austin Smith (Ying et al., 2003; Ying and Smith, 2003). The main advantage of this neural cell differentiation protocol is that differentiation can be easily achieved with adherent monolayer cultures in a defined medium.

Recently, our efforts were focused on the modification of this protocol resulting in a simple procedure for the efficient and reproducible development of neural cells in a comparatively short time. Flow cytometry analysis of a neuron-specific marker protein (microtubule-associated protein 2; MAP2) showed that at as early as day 16 of differentiation the major cell type was the MAP2-positive cell (34% of the total cell population, fig. 3). Studying the influence of selected reference compounds on neural development is currently in progress.

Establishment of serum-free culture medium in the EST

Serum represents an almost universal growth supplement that is effective in most cells. It contains most of the factors required for cell proliferation and maintenance. Using serum-supplemented media therefore reduces the need to spend time developing a specific, optimised medium formulation for every cell type under investigation.

In the EST, for an efficient differentiation of mESC into contracting cardiomyocytes as well as for routine maintenance of the cells, we usually use a cell culture medium supplemented with 15-20% foetal calf serum (FCS). Unfortunately, different serum batches vary considerably with respect to their chemical composition. Certain factors may be deficient in some batches while others may be present at excessive, inhibitory levels. For these reasons serum batches have to be pre-tested for how well they support differentiation and proliferation of mESC before chemical testing in the EST can be performed. In order to further standardise the EST protocol, our aim was to establish

serum-free culture conditions. Chemically defined serum-free culture medium would provide several advantages: (i) improved protocol transfer to other laboratories (ii) improved reproducibility of the differentiation assay, (iii) no interference of undefined serum components with the test substance (reproducible bio-availability) and (iv) application of the EST in automated screening systems.

The effects on proliferation and differentiation of chemically defined serum-free media in combination with selected components have been investigated. Mouse ESC were adapted to various serum-free culture conditions and tested in the validated *in vitro* differentiation assay (Seiler et al., 2006, in press) in comparison to the control using the standard culture medium containing 15% FCS.

Using serum-free Dulbecco's modified Eagle's medium (DMEM) supplemented with specific factors known to promote cardiac development, we were able to exclude FCS completely from the culture medium (Schlechter et al., manuscript in preparation). As shown in figure 4, the differentiation of mESC into contracting cardiomyocytes was as efficient as in the presence of serum.

Conclusions

Reproductive toxicity is one of the most serious side effects a compound can have. Therefore, an important objective for toxicological safety assessment of chemicals and drugs is to evaluate adverse effects on reproduction and embryonic development. The complexity of the reproductive system and the multiple targets for exogenous induction of malformations during embryonic development are the rationale for highly standardised animal experiments, e.g. screening tests or multigeneration studies according to OECD test guidelines.

For reproductive toxicity testing, the estimation of the number of laboratory animals needed under REACH for conventional implementation of the existing OECD guidelines is very high. Thus, testing of these chemicals can hardly be accommodated without the use of predictive cell-based screening systems.

In the field of developmental toxicity testing, the most promising cell-based assay to date is based on ESC of the mouse. As an *in vitro* system, which mirrors both differentiation and growth, the EST is well suited for the evaluation of the teratogenic potential of a compound. Recently, through a number of studies, we were able to improve and expand the validated EST considerably.

New predictive molecular endpoints of differentiation have been effectively implemented in the EST by studying the quantitative expression of marker genes under the influence of the test compound by employing intracellular flow cytometry.

A new differentiation protocol for use in *in vitro* developmental neurotoxicity assays has been developed. Additional major target tissues such as the nervous system may now be included in the test in order to get more precise information on the embryotoxicity of a compound.

Serum-free culture conditions for cell maintenance and differentiation of mESC have been successfully established. Chemical testing can now be performed in a more highly standardised form.

Despite the progress made, there are still some unresolved problems. For regulatory acceptance, the EST requires the enlargement of the database, since the formal validation study focused on pharmaceutical testing. Thus, more industrial chemicals exhibiting different patterns of embryotoxic potential and also a broader range of compounds representing different toxicological mechanisms have to be tested and the database of test chemicals needs to be expanded. Furthermore, in its present form the EST is only applicable for compounds that do not require metabolic activation. However, the metabolic (in-)activation of xenobiotics is an important prerequisite for the accurate determination of the toxic properties of a compound. The introduction of a metabolising system as an adjunct to the vali-

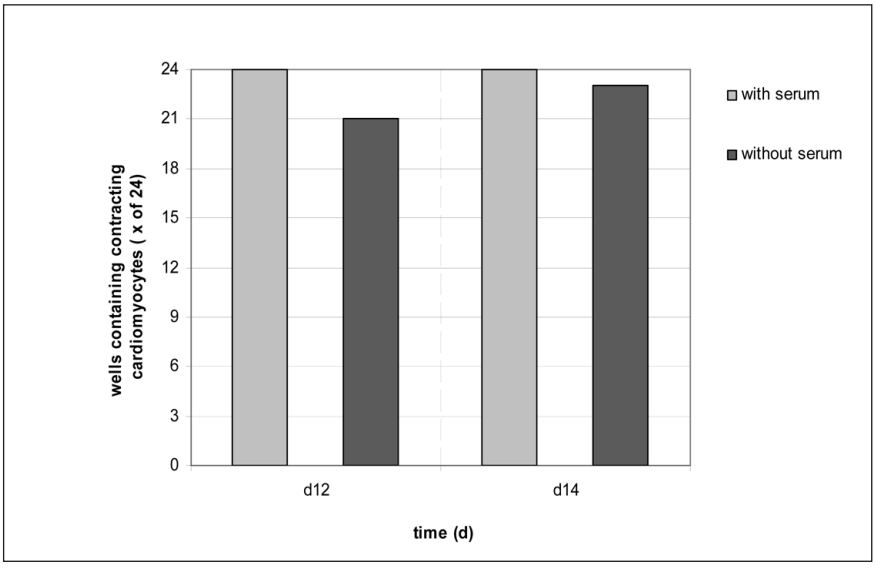


Fig. 4: ESC differentiation assay under serum-free and serum-containing conditions. On day 12 and 14 of the assay, differentiation into contracting myocardial cells was determined under the light microscope. Each well of the 24-well plate was inspected, and the number of wells containing spontaneously contracting cells was recorded. According to the validated EST protocol, the assay is acceptable if at least 21 of the 24 EBs have differentiated into spontaneously contracting myocardial cells. The results from a representative experiment are presented.

dated EST protocol is essential. Finally, to improve the precision of the EST and, in particular, to prevent false negative classifications more cell type-specific endpoints of differentiation (e.g. osteoblasts and chondrocytes) have to be added to the validated EST protocol.

In conclusion, the improved and expanded EST promises to be the first *in vitro* assay for developmental toxicity that may be accepted by regulatory authorities. Furthermore, implementation of the EST in regulatory test guidelines would demonstrate the importance of *in vitro* assays as valuable components of the risk/hazard assessment process.

References

Anon. (2003). Directive 2003/15/EC of the European Parliament and Council directive of 27 February 2003 amending Council directive 76/768/EEC on the approximation of the laws of the Member States relating cosmetic products. *Official J. Eur. Union 2003*, L66, 26-35.

Anon. (2005). The fourth report on the statistics on the number of animals used for experimental and other scientific purposes in the Member States of the European Union. http//www. Ncanl.org/English/Docs/EU/%20stats. Pdf (access 01/2005).

Balls, M. and Hellsten, E. (2002). Statement of the scientific validity of the embryonic stem cell test (EST) – an in vitro test for embryotoxicity. – Statement of the scientific validity of the micromass test – an in vitro test for embryotoxicity. – Statement of the scientific validity of the postimplantation rat whole embryo culture assay – an in vitro test for embryotoxicity. *ATLA 30*, 265-273.

Bain, G., Ray, W. J., Yao, M. and Gottlieb, D. I. (1996). Retinoic acid promotes neuronal and represses mesodermal gene expression in mouse embryonic stem cells in culture. *Biochem. Biophys. Res. Commun. 223*, 691-694.

Boheler, K. R., Czyz, J., Tweedie, D. et al. (2002). Differentiation of pluripotent embryonic stem cells into cardiomyocytes. *Circ. Res. 91*, 189-201.

Buesen, R., Genschow, E., Visan, A. et al. (2005). Establishing predictive molecular markers of differentiation as toxicological endpoints in the embryonic stem cell test (EST), *manuscript in preparation*.

Doetschmann, T., Eistetter, H. R., Schmidt, W. and Kemler, R. (1985). The in vitro development of blastocyst-derived embryonic stem cell lines: Formation of visceral yolk sac, blood islands and myocardium. *J. Embryol. Exp. Morphol. 87*, 7-45.

Eskes, C. and Zuang, V. (2005). Alternative (Non-animal) Methods for Cosmetics Testing: Current Status and Future Prospects. *ATLA 33, Supplement 1*.

Genschow, E., Spielmann, H., Scholz, G. et al. (2002). The ECVAM international validation study on in vitro embryotoxicity tests. Results of the definitive phase and evaluation of prediction models. *ATLA 30*, 151-176.

Genschow, E., Spielmann, H., Scholz, G. et al. (2004). Validation of the embryonic stem cell test (EST) in the ECVAM international validation study on in vitro embryotoxicity. *ATLA 32*, 209-244.

Guan, K., Fürst, D. O. and Wobus, A. M. (1999). Modulation of sarcomere organization during embryonic stem cell-derived cardiomyocyte differentiation. Eur. J. Cell. Biol. 78, 813-23.

Hescheler, J., Fleischmann, B. K., Lentini, S. et al. (1997). Embryonic stem cells: A model to study structural and functional properties in cardiomyogenesis. *Cardiovasc. Res. 36*, 149-162.

Höfer, T., Gerner, I., Gundert-Remy, U., et al. (2004). Animal testing and alternative approaches for the human heath risk assessment under the proposed new European chemicals regulation. *Arch. Toxicol. 78*, 549-564.

ICH (1993), International Conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use. ICH Harmonized Tripartite Guideline: Detection of Toxicity to Human Reproduction for Medicinal Products. Recommended for Adoption at Step 4 of the ICH Process on 24 July 1993 by the ICH Steering Committee. *ICH Secretariat*, Geneva.

Kirschstein, R. and Skirboll, L. R. (2001). Stem cells: Scientific progress and future research directions. Report prepared by the National Institutes of Health. http://www.nih.gov/news/stemcell/scireport.htm, (access 09/2005).

Maltsev, V. A., Wobus, A. M., Rohwedel, J. et al. (1994). Cardiomyocytes differentiated in vitro from embryonic stem cells developmentally express cardiac-specific genes and ionic currents. *Circ. Res. 75*, 233-244.

Martin, G. R., Wiley, L. M. and Damjanov, I. (1977). The development of cystic embryoid bodies in vitro from clonal teratocarcinoma stem cells. *Dev. Biol. 61*, 220-244.

OECD (1983). *OECD Guidelines for the Testing of Chemicals. No. 415*: One-generation Reproduction Toxicity Study, 8 pp. Paris, France: Organization for Economic Cooperation and Development.

OECD (1996). *OECD Guidelines for the Testing of Chemicals. No. 422*: Combined Repeated Dose Toxicity Study with the Reproduction/Developmental Toxicity Screening Test, 14 pp. Paris, France: Organization for Economic Cooperation and Development.

OECD (2001a). *OECD Guidelines for the Testing of Chemicals. No. 414*: Prenatal Developmental Toxicity Study, 6 pp. Paris, France: Organization for Economic Cooperation and Development.

OECD (2001b). *OECD Guidelines for the Testing of Chemicals. No. 416*: Two-generation Reproduction Toxicity Study, 8 pp. Paris, France: Organization for Economic Cooperation and Development.

OECD (2001c). *OECD Guidelines for the Testing of Chemicals. No. 421*: Reproduction/Developmental Toxicity Screening Test, 10 pp. Paris, France: Organization for Economic Cooperation and Development.

Okabe, S., Forsberg-Nilsson, K., Spiro, A. C. et al. (1996). Development of neuronal precursor cells and functional postmitotic neurons from embryonic stem cells in vitro. *Mech. Dev. 59*, 89-102.

Rudnicki, M. A. and Mc Burney, M. W. (1987). Cell culture methods and induction of differentiation of embryonal carcinoma cell lines. In E. J. Robertson (ed), *Teratocarcinoma and Embryonic Stem Cells: A Practical Approach* (19-49). Washington, D.C.: IRL Press.

Sachinidis, A., Fleischmann, B. K., Kolossov, E. et al. (2003). Cardiac specific differentiation of mouse embryonic stem cells. *Cardiovasc. Res. 58*, 278-291.

Schlechter, K., Buesen, R., Spielmann, H. and Seiler, A. (2005). The use of serum-free culture conditions in the embryonic stem cell test: defined culture conditions for growth and differentiation of murine embryonic stem cells, manuscript in preparation.

Seiler, A., Visan, A., Buesen, R. et al. (2004). Improvement of an in vitro stem cell assay (EST) for developmental toxicity by establishing molecular endpoints of tissue-specific development. *Reproductive Toxicology 18*, 231-240.

Seiler, A., Buesen, R., Visan, A. and Spielmann, H. (2006). Use of Murine Embryonic Stem Cells in Embryotoxicity Assays: The Embryonic Stem Cell Test. In: *Embryonic Stem Cells – II*. Totowa, NJ, USA: Humana Press, in press.

Smith, A. G. (2001) Embryo-derived stem cells: Of mice and men. *Annu. Rev. Cell. Dev. Biol. 17*, 435-462.

Spielmann, H., Pohl, I., Döring, B. et al. (1997). The embryonic stem cell test (EST), an in vitro embryotoxicity test using two permanent mouse cell lines: 3T3 fibroblasts and embryonic stem cells. *In Vitro Toxicol. 10*, 119-127.

van der Jagt, K., Munn, S., Torslov, J. and de Bruijin,J. (2004). Alternative approaches can reduce the use of test animals under REACH. http://ihcp.jrc.cec.eu.int/DOCUMENTATION/ECB/2004/Reducing_the_use_of_test_animals_under_REACH_IHCP_report.pdf (access 09/2005).

Williams, R. L., Hilton, D. J., Pease, S. et al. (1988). Myeloid leucemia inhibitory factor maintains the developmental potential of embryonic stem cells. *Nature 336*, 684-687.

Ying, Q.-L., Stavridis, M., Griffiths, D. et al. (2003). Conversion of embryonic stem cells into neuroectodermal precursors in adherent monocultures. *Nature Biotechnology 21*, 183-186.

Ying, Q.-L. and Smith, A. G. (2003). Defined Conditions for neural commitment and differentiation. *Methods in Enzymology 365*, 327-349.

Acknowledgements

We gratefully acknowledge funding of the projects by the German Federal Ministry for Education and Research (BMBF, grant no. 0312312) as well as the support of the FP6 Project ReProTect (LSHB-CT-2004-503257 ReProTect).

Correspondence to

Dr. Andrea Seiler
ZEBET at the BfR
Diedersdorfer Weg 1
12277 Berlin
Germany
e-mail: seiler.zebet@bfr.bund.de

In Vitro Multipotency of Human Bone Marrow (Mesenchymal) Stem Cells

Sarah Snykers[1], Tamara Vanhaecke[1], Peggy Papeleu[1], Tom Henkens[1], Mathieu Vinken[1], Greetje Elaut[1], Ivan Van Riet[2] and Vera Rogiers[1]

[1]Dept. Toxicol., Vrije Universiteit Brussel, Brussels, Belgium, [2]Dept. Med. Oncol.-Hematol.-Stem Cell Lab., Academic. Hospital, Vrije Universiteit Brussel, Brussels, Belgium

Summary

The capability of mesenchymal stem cells (MSC) to differentiate in vitro into hepatocytes by exposure to liver-specific cytokines was investigated. Simultaneous exposure of MSC to a mixture of hepatogenic cytokines only stimulated neuroectodermal and mesodermal differentiation. However, sequential exposure, resembling the order of secretion during liver embryogenesis, induced both glycogen storage and expression of cytokeratin 18. In order to trigger further endodermal differentiation, cells were exposed to trichostatin A (TSA). The latter up-regulated albumin secretion, a typical functional property of primary hepatocytes.

In conclusion, MSC acquire trilineage potential under "hepatocyte-specific conditions". TSA improves differentiation of MSC towards hepatocyte-like cells.

Keywords: bone marrow stem cells, hepatocytes, embryonic development, histone-deacetylase inhibitor, in vitro

Introduction

The liver is the principal organ for xenobiotic biotransformation and thus a key target for drug-induced toxicity (Gibson and Skett, 1994). Isolated primary hepatocytes and their cultures are therefore widely used in preclinical pharmaco-toxicological research and testing. A problem encountered, however, is the relative short lifespan and the rapid decline of liver-specific functions as a function of culture time (LeCluysse et al., 1995; Rogiers et al., 1995; Papeleu et al., 2002). Attempts have been made to cultivate functional hepatocytes for longer periods by using soluble medium components and by affecting cell-cell and cell-matrix interactions (Guillouzo, 1998; Rogiers et al., 1995). The ideal *in vitro* model, however, does not yet exist. Different alternatives have been explored to overcome these difficulties. In recent years, adult-derived stem cells became a hot topic in the field of molecular and cellular biology, for clinical application as well as for pharmaco-toxicological purposes. Indeed, stem cells are clonogenic and capable of both self-renewal and multilineage differentiation (Krause et al., 2001; Huttmann et al., 2003). *In vivo* as well as *in vitro* studies have provided evidence that stem cells can overcome germ lineage restrictions and express molecular characteristics of cells of different tissue origin (Krause et al., 2001; Huttmann et al., 2003).

The best-characterised stem cell compartment is the bone marrow, which consists of two stem cell populations, referred to as haematopoietic and mesenchymal stem cells (Huttmann et al., 2003). The latter were first described by Friedenstein et al. as a population of cells isolated from the bone marrow and capable of differentiation into bone, adipocytes, chondrocytes, osteoblasts, osteoprogenitors, skeletal myocytes, tendon and bone marrow stromal cells (Friedenstein, et al., 1976, 1987).

Schwartz et al. described a population of cells in postnatal rat bone marrow named multipotent adult progenitor cells (MAPC) that were not only able to differentiate into most mesodermal cell types, but also into neuroectodermal and endodermal (hepatocytes) cell types (Schwartz et al., 2002).

In the present study, we investigated whether human MSC also have the potential to differentiate *in vitro* into endodermal cells such as hepatocytes. Experiments were undertaken in the presence of liver-specific cytokines, added either as a cocktail [(fibroblast growth factor-4 (FGF-4), hepatocyte growth factor (HGF), insulin-transferrin-sodium-selenite (ITS) and dexamethasone)] or sequentially, in a manner that closely reflects their temporal expression during *in vivo* hepatogenesis (FGF-4, followed by HGF, followed by a combination of HGF, ITS and dexamethasone).

Materials and methods

Isolation and expansion of MSC

Expanded MSC were obtained from the Academic Hospital, Vrije Universiteit Brussel. MSC were collected from 5-30 ml aspirates from the sternum of haematologically healthy donors and expanded *in vitro* at low density for 4 passages in medium containing foetal calf serum (Friedenstein, et al., 1976, 1987).

Multilineage differentiation of MSC

MSC (CD45-, Thy^+) were cultivated at 21.5 x 10^3 cells / cm^2 on 1 mg/ml collagen gel type I in the presence of liver-specific cytokines, added either as a cocktail [basal medium (Jiang et al., 2003) + 10 ng/ml FGF-4, 20 ng/ml HGF (both from R&D Systems), 1x ITS and 20 µg/l dexamethasone (both from

Sigma)] or sequentially (days 0-3: basal medium + 10 ng/ml FGF-4; days 3-6: basal medium + 20 ng/ml HGF; from day 6 on: basal medium + 20 ng/ml HGF + 1x ITS and 20 µg/l dexamethasone). Differentiation media were changed every 3 days. 1 µM TSA (Sigma) was added from day 6 on.

Cytological staining

Cells were fixed with 10% formalin for 10 min at room temperature (neuroectodermal, adipogenic, hepatocyte differentiation) or with MeOH for 2 min at -20°C (osteogenic differentiation). After fixation, nerve vessels were identified by Bodian staining (Wullimann et al., 1999). Adipocytes were identified by red lipid droplets after staining with Sudan III (Reyes et al., 2001). Mineralised nodules were stained black with the von Kossa technique, whereas unmineralised nodules were stained yellow (Reyes et al., 2001). Periodic-acid-Schiff (PAS) staining was used to determine glycogen storage, a functional parameter of endogenic differentiation (Schwartz et al., 2002). As control, PAS staining was performed in the presence of amyloglucosidase (Sigma).

Immunofluorescence

Immunostaining was performed as previously described (Schwartz et al., 2002).

Albumin ELISA

Albumin concentrations, secreted into the culture media, were quantified by ELISA (Koebe et al., 1994).

Statistics

Results are expressed as means ± SD of five independent experiments. Statistical analyses were performed using one-way ANOVA followed by Student's t-test. The significance level was set at 0.05.

Results

Trilineage differentiation of low-density MSC cultures

To evaluate whether MSC had trilineage potential like MAPC, specific cytological staining was performed.

Adipogenic, osteogenic and neurogenic differentiation

Undifferentiated MSC did not stain positive for any of the cytological markers.

Simultaneous exposure of MSC to a mixture of hepatogenic cytokines resulted in a heterogeneous population of cells (fig. 1). More specifically, adipogenic differentiation was seen from day 5 on (fig. 1B). After 17 days of cultivation, more than 80% of the cells were loaded with lipid droplets and stained red with Sudan III (fig. 1D). In addition, the presence of calcium phosphate, characteristic for osteoblasts, could be shown from day 5 on by von Kossa staining (fig. 1F). Bodian staining clearly revealed the presence of nerve fibres (black staining) on day 8 (fig. 1G).

In order to obtain more homogeneous MSC-derived cell cultures, MSC were separately treated with each of the cytokines. It appeared that only adipocytes formed in the pres-

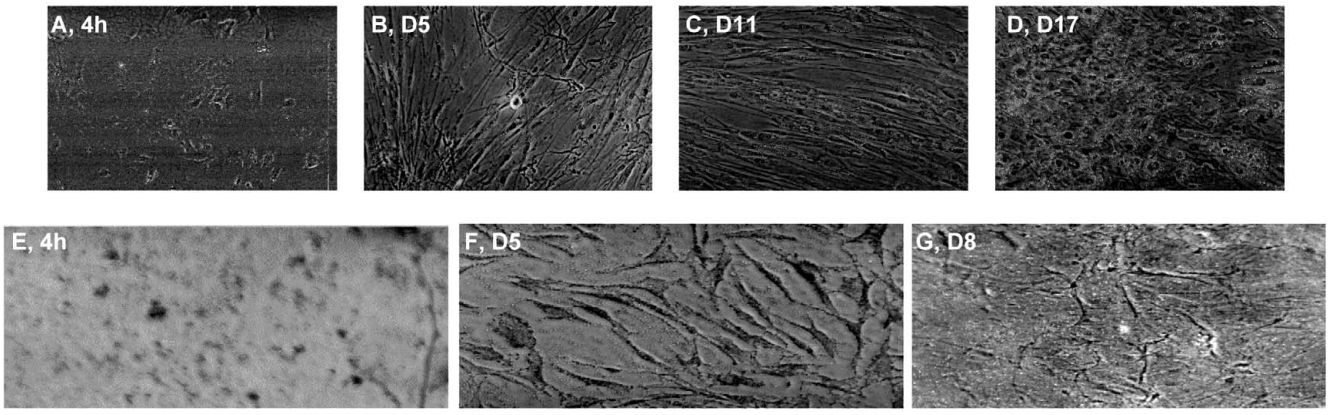


Fig. 1: Adipogenic, osteogenic and neuroectodermal differentiation was demonstrated by means of Sudan III (A-D), Von Kossa (E,F) and Bodian (G) staining, respectively. MSC were simultaneously exposed to all liver-specific cytokines. 10 x 10 magnification.

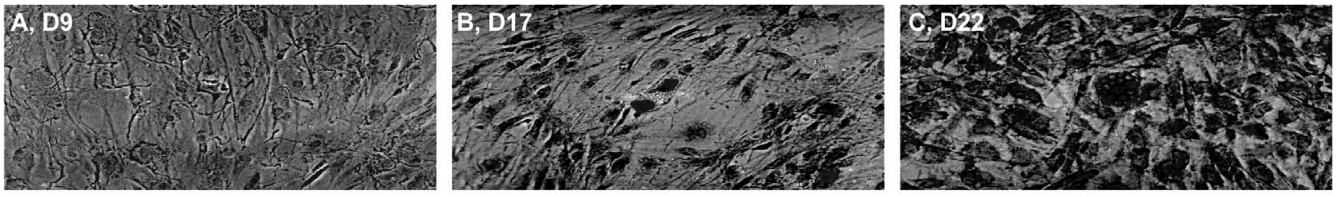


Fig. 2: PAS-staining showed glycogen storage (A-C) from day 9 on. MSC were sequentially treated with liver-specific cytokines for 9, 17 and 22 days, respectively. 10 x 10 magnification.

ence of ITS and dexamethasone, whereas FGF-4 induced the appearance of osteoblasts and neuroectodermal-like cells (data not shown).

Endogenic differentiation

Since no cells with typical morphological characteristics of hepatocytes were seen upon treatment with the cocktail of liver-specific cytokines, MSC were sequentially exposed to FGF-4, HGF, ITS and dexamethasone in a time-dependent order that closely resembles their secretion pattern during *in vivo* liver ontogeny. Under these conditions, glycogen uptake was seen from day 9 on (fig. 2A) and was clearly up-regulated throughout the culture time (fig. 2B-C). Moreover, cells stained positive for cytokeratin 18 (fig. 3A), a cytoskeletal filament present in hepatocytes, while control immunoglobulins did not (fig. 3B). Nevertheless, the differentiated cells did not resemble hepatocytes morphologically.

TSA, a trigger for endodermal differentiation of hMSC

In an attempt to enhance endogenic differentiation of MSC, TSA, a selective and reversible histone-deacetylase (HDAC) inhibitor (Xu, et al., 1997) was added to the culture medium.

Hepatic functionality

The secretion of albumin, a typical functional property of primary hepatocytes, was taken as a first criterion for hepatic functionality of differentiated MSC.

MSC were sequentially exposed to liver-specific cytokines, as during *in vivo* hepatogenesis. 1 µM TSA was added from day 6 on, corresponding to promising culture conditions used for hepatic differentiation of MAPC (unpublished data). Interestingly, addition of 1 µM TSA to sequentially-treated MSC induced a significant upregulation of albumin secretion from day 15 on ($p<0.01$, one-way ANOVA and Student's t-test) compared with control cultures (sequential) (fig. 4). At day 17, albumin secre-

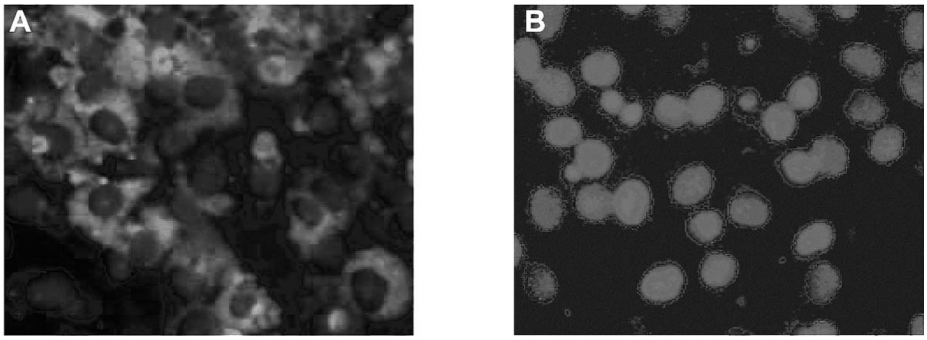


Fig. 3: Immunofluorescence was performed for cytokeratin18-FITC at day 7 (A), compared to mouse immunoglobuline (B) as negative control. MSC were sequentially treated with liver-specific cytokines. 32 x 10 magnification.

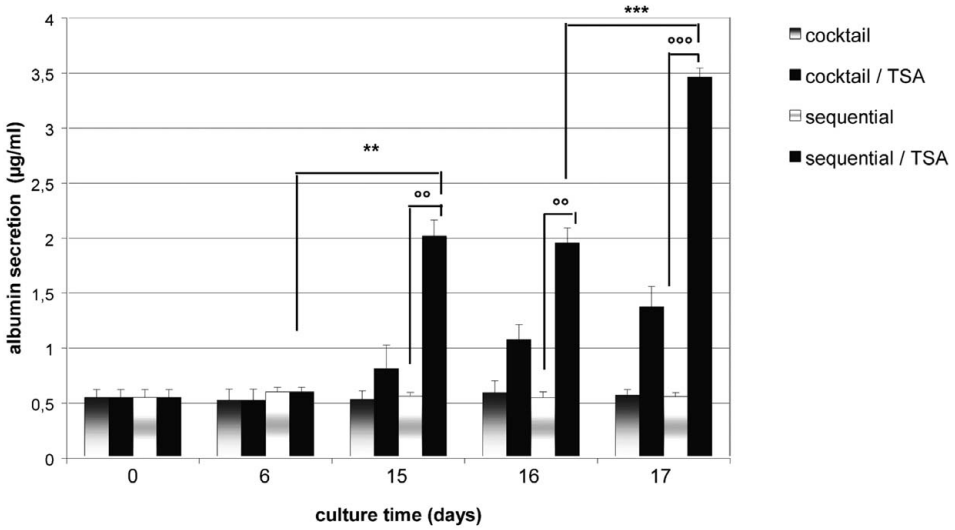


Fig. 4: Albumin secretion into the media was measured by ELISA. MSC were either sequentially or simultaneously exposed to liver-specific cytokines. 1 µM TSA was added from day 6 on. The graph is representative for five independent experiments.
, *: Albumin-secretion of MSC cultivated in sequential + TSA is significantly upregulated during culture time with $p<0.01$ and $p<0.001$ (Oneway Anova + t-test).
°°, °°°: Albumin-secretion of MSC cultivated in sequential + TSA is significant higher than control cultures (sequential) with $p<0.01$and $p<0.001$ (Oneway Anova + t-test).

tion even reached comparable levels with those observed in 2-day old monolayer cultures of primary rat hepatocytes (Vanhaecke et al., 2004). Up-regulation of albumin secretion was also seen upon exposure to a cytokine-cocktail supplemented with 1 µM TSA. However, this result was not statistically significant and was less distinctive than that observed after sequential treatment. MSC cultivated without TSA did not secrete albumin (fig. 4).

Morphological features

This inductive effect of TSA on the differentiation process of MSC to hepatocyte-like cells was well supported by microscopic analysis of the cell morphology. Indeed, upon exposure to TSA, a complete transformation of the cell morphology could be observed, regardless of the experimental set-up (fig. 5). After 3-5 days TSA-treated MSC formed epithelioid cells with clear, round nuclei. Fibroblastic cells, however, persisted throughout the culture time, particularly upon exposure to the cytokine-cocktail (fig. 5A2).

Discussion and conclusions

Hepatocyte-based *in vitro* models are important tools for pharmaco-toxicological research and regulatory testing of xenobiotics (Gibson et al., 1994). A serious drawback, however, is their limited lifespan and rapid loss of differentiated hepatic properties (LeCluysse et al., 1995; Rogiers et al., 1995; Papeleu et al., 2002). The use of stem cells could offer a solution. It is well recognised that in adult mammals, including humans, a number of tissues are continuously regenerated from immature cells (i.e. skin, intestinal epithelia, blood cells, the olfactory bulb in the brain) (Krause et al., 2001; Huttmann et al., 2003; Friedenstein et al., 1976; Friedenstein et al., 1987; Schwartz et al., 2002; Jiang et al., 2003; Reyes et al., 2001). The interest in adult stem cells has in particular been triggered by the numerous ethical dilemmas surrounding the use of embryonic stem cells in pre-clinical and clinical research (Henningson et al., 2003). Therefore, the ability to isolate, cultivate and manipulate multipotent stem cells from non-embryonic origin would provide researchers with an unlimited source for cell and organ development studies, pharmaco-toxicological research and regulatory testing.

In the present study, we investigated whether MSC, like MAPC, are able to differentiate *in vitro* into cell types of all three germ layers. In particular, the endodermal differentiation received most attention. MSC were hence exposed to well-defined hepatogenic cytokines, added either as a cocktail or sequentially, in a time dependent order as seen during liver embryogenesis.

Upon exposure to a cocktail of liver-specific cytokines, low-density MSC differentiated into a heterogeneous population of

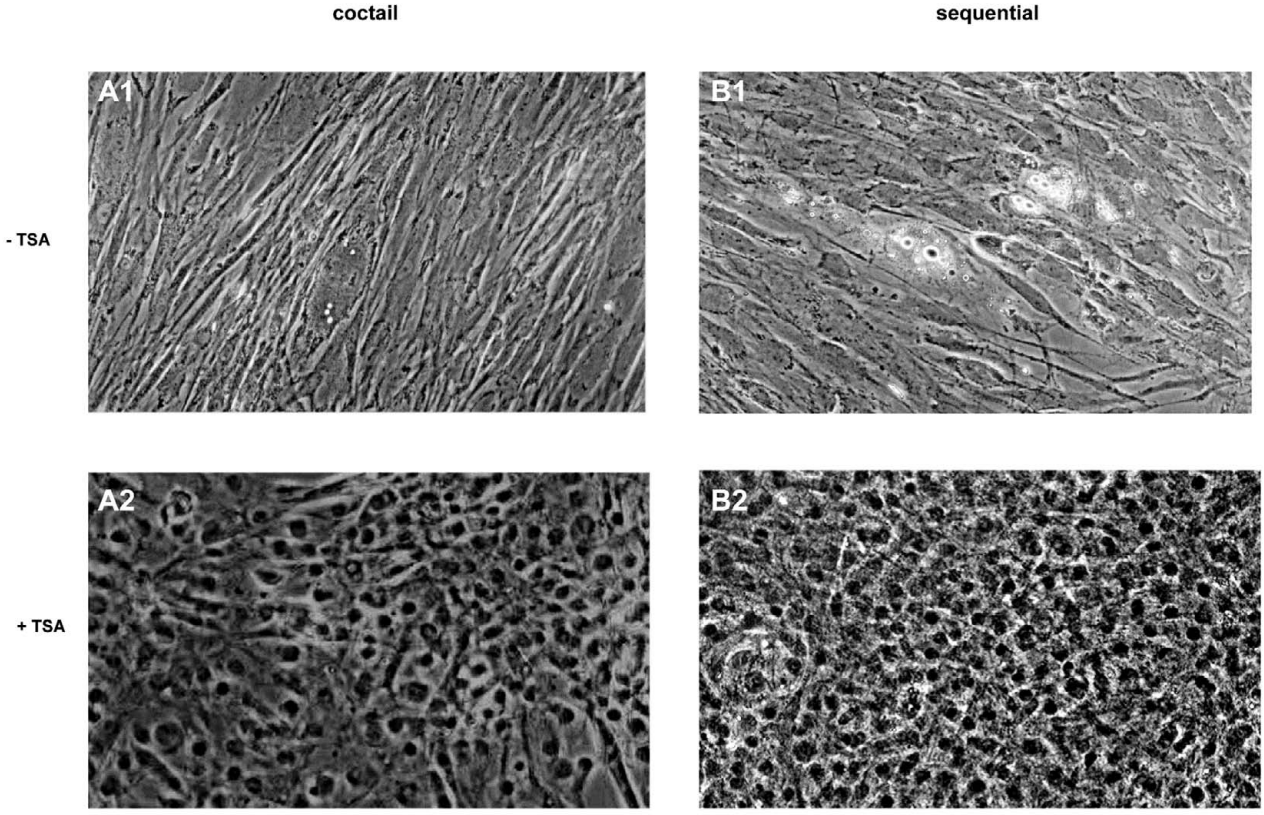


Fig. 5: Light-microscopic analysis of 17-day old differentiated MSC upon simultaneous (A) or sequential (B) exposure to liver-specific cytokines in the absence (1) and presence (2) of 1 µM TSA, respectively; 20 x 10 magnification, phase contrast.

mesodermal (adipocytes and osteoblasts) and ectodermal (neuroectodermal-like cells) cells. Although lipid droplets (adipogenic differentiation) also appear in hepatic epithelium, the differentiated cells did not display an epithelial-like morphology. They neither expressed other hepatocyte-specific markers such as alphafoetoprotein, albumin or cytokeratin 18, nor did they take up glycogen (data not shown).

Conversely, sequential exposure of MSC to hepatogenic cytokines induced not only expression of the mid-late hepatic marker cytokeratin 18 from day 7 on, but also caused an upregulation of glycogen storage from days 9 to 22. However, cells did not morphologically resemble adult hepatocytes. As a trigger for further differentiation of MSC towards endodermal lineage, TSA, a drug candidate for hyperproliferative disorders, was introduced into the present culture media. TSA causes an hyperacetylation of histone proteins, resulting in an increased accessibility of target DNA for transcription factors, and thus also liver-enriched transcription factors, thereby facilitating transcription of the target genes (Marks et al., 2001; Yoshida et al., 1990). As such, TSA and HDAC-inhibitors in general, have been shown to induce differentiation, apoptosis and cell cycle arrest in tumour cells (Yamashita et al., 2003). In primary cells, e.g. cultured hepatocytes, it was found that 1 μM TSA induced cell cycle arrest during G0/G1 and G1/S phase in EGF-stimulated cells but it did not induce apoptosis (Papeleu et al., 2003). These findings triggered us to add TSA to the culture media in order to positively affect the differentiation process towards the endodermal lineage. Indeed, upon sequential exposure to liver-specific cytokines, TSA caused a significant upregulation of albumin secretion, a typical functional property of hepatocytes, to levels similar to those found in monolayer cultures of 2-day old primary rat hepatocytes (Vanhaecke et al., 2004). Furthermore, microscopic analysis supported these observations. TSA-treated MSC adopted an epithelial-like morphology, regardless of the culture conditions. However, a higher number of fibroblastic cells persisted throughout the culture time upon exposure to the cytokine cocktail.

Currently, the culture conditions for advanced differentiation of MSC into mature hepatocytes are being optimised.

References

Friedenstein, A. J., Chailakhyan, R. K. and Gerasimov, U. V. (1987). Bone marrow osteogenic stem cells: in vitro cultivation and transplantation in diffusion chambers. *Cell Tissue Kinet. 20(3)*, 263-272.

Friedenstein, A. J., Gorskaja, J. F. and Kulagina, N. N. (1976). Fibroblast precursors in normal and irradiated mouse hematopoietic organs. *Exp. Hematol. 4(5)*, 267-274.

Gibson, G. G. and Skett, P. (1994). *Introduction to drug metabolism*, second edition. Glasgow: Blackie Academic & Professional, Chapman & Hall.

Guillouzo, A. (1998). Liver cell models in in vitro toxicology. *Environ. Health Perspect. 106*, 511-532.

Henningson, C. T., Stanislaus, M. A. and Gewirtz, A. M. (2003). Embryonic and adult stem cell therapy. *J. Allergy Clin. Immunol. 111*, 745-753.

Huttmann, A., Li, C. L. and Duhrsen, U. (2003). Bone marrow-derived stem cells and plasticity. *Ann. Hematol. 82*, 599-604.

Jiang, Y., Henderson, D., Blackstad, M. et al. (2003). Neuroectodermal differentiation from mouse multipotent adult progenitor cells. *PNAS 100*, 11854-11860.

Koebe, H. G., Wick, M., Cramer, U. et al. (1994). Collagen gel immobilisation provides a suitable cell matrix for long term human hepatocyte cultures in hybrid reactors. *Int. J. Artif. Organs. 17(2)*, 95-106.

Krause, D. S., Theise, N. D. and Collector, M. I. (2001). Multi-organ, multi-lineage engraftment by a single bone marrow-derived stem cell. *Cell 105*, 369.

LeCluysse, E. L., Bullock, P. L. and Parkinson, A. (1995). Strategies for restoration and maintenance of normal hepatic structure and function in long-term cultures of rat hepatocytes. *Adv. Drug Del. Rev. 22*, 133-186.

Marks, P. A., Richon, V. M., Breslow, R. and Rifkind, R. A. (2001). Histone deacetylase inhibitors as new cancer drugs. *Curr. Opin. Oncol. 13(6)*, 477-483.

Papeleu, P., Elaut, G., Rogiers, V. and Vanhaecke, T. (2002). Cell cultures as in vitro tools for biotransformation studies. In Recent Research Developments in Drug Metabolism and Disposition, Pandalai SG (ed.), *Transworld Research Network* (1, 199-234). Kerala, India.

Papeleu, P., Loyer, P., Vanhaecke, T. et al. (2003). Trichostatin A induces differential cell cycle arrests but does not induce apoptosis in primary cultures of mitogen-stimulated rat hepatocytes. *J. Hepatol. 39(3)*, 374-82.

Proposal for a directive of the European Parliament and of the Council amending Council Directive 67/548/EEC in order to adapt it to Regulation (EC) of the European Parliament and of the Council concerning the registration, evaluation, authorisation and restriction of chemicals (SEC(2003 1171)/*COM/ 2003/0644 final –COD 2003/0257*/, October29, 2003.

Reyes, M., Lund, T., Lenvik, T. et al. (2001). Purification and ex vivo expansion of postnatal human marrow mesodermal progenitor cells. *Blood 98(9)*, 2615-2625.

Rogiers, V., Blaauboer, B., Maurel, P. et al. (1995). Hepatocyte-based in vitro models and their application in pharmaco-toxicology. Toxicol. *In Vitro 9*, 685-694.

Schwartz, R. E., Reyes, M., Koodie, L. et al. (2002). Multipotent adult progenitor cells from bone marrow differentiate into functional hepatocyte-like cells. *J. Clin. Invest. 109(10)*, 1291-1302.

Vanhaecke, T., Henkens, T., Kass, G. E. and Rogiers, V. (2004). Effect of the histone deacetylase inhibitor trichostatin A on spontaneous apoptosis in various types of adult rat hepatocyte cultures. *Biochem. Pharmacol. 68(4)*, 753-760.

Wullimann, M. F. and Puelles, L. (1999). Postembryonic neural proliferation in the zebrafish forebrain and its relationship to prosomeric domains. *Anat. Embryol. 199*, 329-348.

Xu, L., Ruh, T. S. and Ruh M. F. (1997). Effect of the histone deacetylase inhibitor trichostatin A on the responsiveness of rat hepatocytes to dioxin. *Biochem. Pharmacol. 53(7)*, 951-957.

Yamashita, Y., Shimada, Y., Harimoto, N. et al. (2003). Histone deacetylase inhibitor trichostatin A induces cell-cycle arrest/apoptosis and hepatocyte differentiation in human hepatoma cells. *Int. J. Cancer. 103(5)*, 572-576.

Yoshida, M., Kijima, M., Akita, M. and Beppu, T. (1990). Potent and specific inhibition of mammalian histone deacetylase both in vivo and in vitro by trichostatin A. *J. Biol. Chem. 265(28)*, 17174-17179.

Acknowledgements

Research was funded by a Ph.D. grant of the Institute for the Promotion of Innovation through Science and Technology in Flanders (IWT-Vlaanderen) and grants from the Research Council (OZR) of the Vrije Universiteit Brussel, Belgium.
The authors thank Prof. Y. Vander Heyden and B. Dejaegher for their valuable statistical advice and G. De Pauw, E. Desmedt, B. Degreef, H. Mertens and S. Coppens for their dedicated technical assistance.

Abbreviations:

fibroblast growth factor-4 (FGF-4);
hepatocyte growth factor (HGF);
histone-deacetylase (HDAC);
insulin-transferrin-sodium-selenite (ITS);
mesenchymal stem cells (MSC);
multipotent adult progenitor cells (MAPC);
Periodic-acid-Schiff (PAS) staining;
trichostatin A (TSA).

Correspondence to

Sarah Snykers
Dept. Toxicology
Vrije Universiteit Brussel
Laarbeeklaan 103
1090 Brussels
Belgium
e-mail: Sarah.Snykers@vub.ac.be.

Session 7.2
Innovative approaches for alternative methods development

RNA Interference: A Novel Alternative Approach in Nephrotoxicity Studies

Eric Campbell, Craig Slattery, Niamh Tuite, Stephen Nolan, Martin Leonard, Michael P. Ryan and Tara McMorrow

Department of Pharmacology, Conway Institute, School of Biomolecular and Biomedical Research, University College Dublin, Dublin, Ireland

Summary

One of the main goals in functional genomics has been the development of tools that allow easy manipulation of gene expression levels that would be suitable for high throughput screening. RNA interference (RNAi) has emerged as one of the preferred approaches to achieve this goal. It is an important biological mechanism in the regulation of gene expression in animals and plants.

Here we evaluated the use of RNAi for knockdown of specific gene expression as an alternative to the production of transgenic animals. RNAi development may be in the early stages, however, the real and theoretical advantages of this system in reducing the use of animals merit further investigations.

Keywords: RNA interference, siRNA, nephrotoxicity, EMT, Rap, renal fibrosis

Introduction

RNA interference (RNAi) is a term used to describe the sequence-specific suppression of genes, initiated by double-stranded RNA (dsRNA) oligonucleotides homologous to the sequence of the silenced gene (Hammond et al., 2001). RNAi was first observed in *Caenorhabditis elegans* as a response to exogenously introduced long double-stranded RNA (dsRNA) (Fire et al., 1998). While the mechanism of RNAi has yet to be fully characterised, silencing appears to be mediated by duplexes of 21-23 nucleotides (nt), called small interfering RNAs (siRNAs), which are degradation products of longer double-stranded RNAs, cleaved by the RNAse III-like enzyme Dicer. The resulting siRNAs are incorporated into a multi-component complex known as RISC (RNA-induced silencing complex) that targets and cleaves mRNA complementary to the siRNAs (Hamilton et al., 1999; Tuschl et al., 1999; Zamore et al., 2000; Bernstein et al., 2001; Elbashir et al., 2001; Hammond et al., 2001). An overview of the RNAi mechanism is shown in fig. 1.

More recently, short endogenous RNAs sharing similarity to siRNAs, called microRNAs (miRNAs), have been identified as key post-transcriptional regulators in such varied organisms as plants, nematodes and mammals (Lagos-Quintana et al., 2001; Lau et al., 2001; Lee et al., 2001; Reinhart et al., 2002). These miRNAs are initially transcribed as a long RNA and then processed to a pre-miRNA of ~70 nt (Lee et al., 2002). This pre-miRNA forms an imperfect hairpin structure, which is then processed by Dicer enzyme to produce the mature, single-strand ~21-23 nt miRNA (Grishok et al., 2001; Hutvagner et al., 2001; Ketting et al., 2001).

While this is an innate process in cells, most probably an evolutionary mechanism to protect against cellular attack from invasive viral RNAs, the mechanism provides researchers with

an excellent tool for gene silencing. Knock-out mice are traditionally the ultimate way of demonstrating gene function, however knocking out a gene that is crucial for development can be a terrible waste if the mice die before a discernable phenotype can be established. RNAi allows researchers to knock out several genes *in vitro* in a fraction of the time required to generate a knock-out mouse.

3R relevance

The relevance of this article to the goal of replacement, refinement and reduction of animal testing is due to the future ability of RNAi to possibly replace most if not all animals used in the generation of knock-out mice.

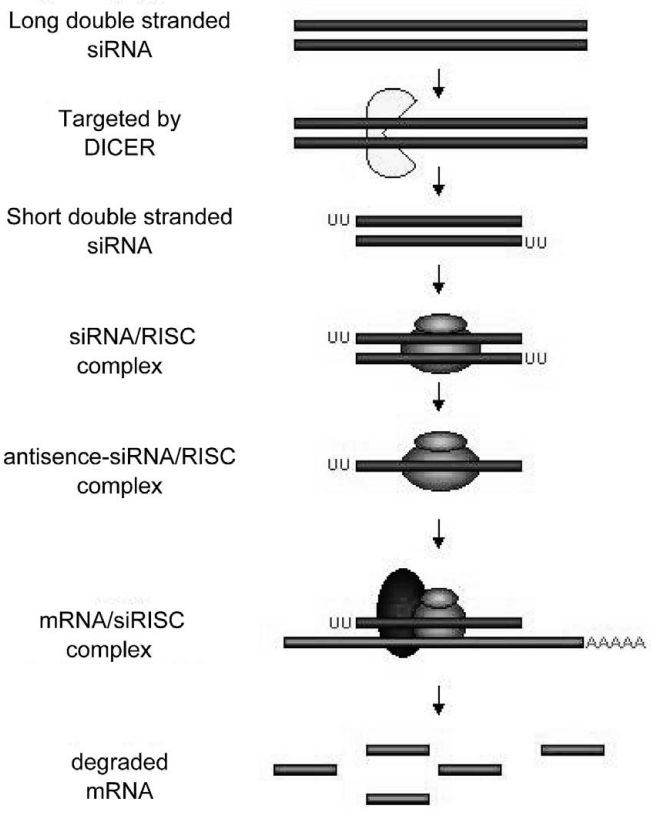


Fig. 1: Hypothesised mechanism of RNAi-mediated gene silencing
Long dsRNA molecules are targeted by the DICER enzyme and cleaved into smaller dsRNA molecules. These small dsRNAs are bound by the RISC complex, which separates and discards the non-complementary strand. The complementary strand directs the RISC complex to the target mRNA and cleaves it. Cleaved RNA is degraded by cellular proteases.

Materials and methods

Reagents

Silencer™ siRNA Construction Kit was obtained from Ambion Inc. Lipofectamine 2000 transfection kit was obtained from Invitrogen. siPORT amine, siPORT lipid transfection kits were obtained from Ambion Inc. Antibodies were obtained from BD Biosciences. PCR primers were supplied by Sigma-Genosys (UK). Dulbecco's Modified Eagles Medium / Nutrient Mix F12 (DMEM/F12) was obtained from Gibco BRL Ltd (UK). Cell culture dishes and slides were purchased from Falcon (UK).

Cell culture

HK-2 cells were maintained in Dulbecco's Modified Eagles Medium/Nutrient Mix F12 (DMEM/F12) (Gibco BRL Ltd, UK) containing 5 µg/ml insulin, 5 µg/ml transferrin, 5 ng/ml selenium, 36 ng/ml hydrocortisone, 4 pg/ml triiodo-l-thyronine, 10 ng/ml epidermal growth factor, 50 U/ml penicillin, 50 µg/ml streptomycin and 2mM glutamine. Culture medium was changed every second day. Cells were grown to confluency in 75 cm^2 Falcon flasks and maintained at 37ºC in a humidified atmosphere containing 95% air, 5% CO_2.

siRNA design

Sequences targeted for RNAi were scanned for occurrences of AA, using Target Finder (http://www.ambion.com/techlib/misc/siRNA_finder.html). This online tool identifies occurrences of AA in a target sequence and lists it with the subsequent 19 downstream nucleotides. Target siRNAs that had a low GC content (30-50%) and, where possible, were not within 50 bases of the ATG start codon were considered. Targets were analysed by BLAST search and only siRNAs that had no significant homology with other genes were selected.

siRNA synthesis and transfection

Small interfering RNAs (siRNAs) were synthesised with the Silencer™ siRNA Construction Kit (Ambion) as directed by the manufacturer, with the exception that half reaction steps were used. Template siRNA oligonucleotides were designed and DNA oligonucleotides corresponding to the sense and antisense strands of the siRNAs were synthesised and desalted (Sigma-Genosys). In separate reactions the sense and antisense DNA oligos were annealed to the T7 promoter primer and subsequently filled in with the Klenow fragment of DNA polymerase. RNA was transcribed from these double-stranded DNA templates using T7 RNA polymerase. The sense and antisense RNA reactions were hybridised and digested with RNase to leave double-stranded 21-mer RNA duplexes with UU overhangs. siRNAs were purified by column chromatography and quantified by UV spectrophotometry. siRNA corresponding to 30 nM was mixed with Lipofectamine 2000 (Invitrogen) as directed by the manufacturer and the mixture was used for transfection of the cells.

RT-PCR analyses

Total RNA was extracted from cells using TRIzol reagent (Life Technologies-BRL). Briefly, a 5 µg sample of total RNA from each treatment group was used to generate cDNA with the

Gibco-BRL RT kit. 2 µl of the cDNA was taken from each sample and used in a PCR reaction containing 1 µl of forward primer (100 ng/µl), 1 µl of reverse primer (100 ng/µl), 2 µl of 10 mM dNTPs, 5 µl of 10X buffer, 3 µl of 25 mM MgCl2, 0.25 µl of Taq Polymerase (Promega). The number of PCR cycles performed was determined to be within the linear range of the reactions.

Western blot analyses

Whole cell extracts of control and treated cells were obtained in radioimmune precipitation assay buffer (50mM Tris-HCl, 150mM NaCl, pH 7.5), and analysed for the indicated molecules by Western blot and enhanced chemiluminescence detection as described previously (Slattery et al., 2005).

Results

RNAi optimisation

To optimise RNAi in HK-2 cells, siRNAs directed against GAPDH were synthesised and delivered using various transfection reagents (siPORT amine, siPORT lipid, Lipofectamine 2000) and concentrations (1nM-100 nM). Initially, higher concentrations (50 and 100 nM) of siRNA were used to identify the most suitable transfection reagent. RT-PCR was performed on RNA isolated from HK-2 cells transfected with 50 or 100 nM siRNA using 3 different transfection reagents. It was determined that Lipofectamine 2000 was the most effective transfection reagent for siRNA transfection of HK-2 cells. However high concentrations (50-100 nM) of siRNA can mediate non-specific gene silencing (Elbashir et al., 2001), therefore it was necessary to further optimise conditions using lower concentrations (10 or 30 nM) of siRNA with Lipofectamine 2000. Using RT-PCR for GAPDH, 30 nM of GAPDH siRNA was established to have approximately 50% silencing efficiency. β-Actin was used as a loading control. A similar decrease was detectable by Western blot analysis.

Specific gene silencing using RNA interference

Rap genes were identified as being differentially expressed in *in vitro* models of nephrotoxicity, therefore Rap protein expression was silenced using gene-specific siRNAs. In an *in vitro* model of nephrotoxicity, siRNAs directed against either Rap1 or Rap2 were investigated. In treated cells transfected with Rap1 siRNA there was a decrease in Rap1 protein levels, but not in Rap2 levels. Concurrently, TECs treated with Rap2 siRNA demonstrated a decrease in Rap2, but not Rap1 protein levels.

Discussion and conclusion

Functional investigation of the complex regulation of molecular switches and their effectors are key to understanding organ toxicity, including nephrotoxicity. Nephrotoxicity is caused by several drugs, such as immunosuppressive agents, and is associated with the development of renal fibrosis. We have established several *in vitro* models of nephrotoxicity and have identified key genes, using both microarray and differential gene expression (SSH) technology, which we believe are involved in this process. We have used RNAi to analyse the role of these genes in the development of nephrotoxicity and we are currently examining the possible involvement of endogenous miRNA regulation of candidate genes. A number of interesting differentially expressed genes are being examined further as potential therapeutic targets.

In conclusion we have demonstrated that silencing of key genes with RNAi has helped to elucidate their role in the development of renal fibrosis and nephrotoxicity. These results not only indicate the specificity and sensitivity of the siRNAs used, but also show the potential of RNAi as a replacement for animal studies. Finally, we believe that the potential role of RNAi as an alternative to animal models is just beginning to be realised.

References

Elbashir, S. M., Harborth, J. et al. (2001). Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. *Nature 411*, 494-498.

Fire, A., Xu, S. et al. (1998). Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. *Nature 391*, 806-811.

Grishok, A., Pasquinelli, A. E. et al. (2001). Genes and mechanisms related to RNA interference regulate expression of the small temporal RNAs that control C. elegans developmental timing. *Cell 106*, 23-34.

Hamilton, A. J. and Baulcombe, D. C. (1999). A species of small antisense RNA in posttranscriptional gene silencing in plants. *Science 286*, 950-952.

Hammond, S. M., Caudy, A. A. et al. (2001). Post-transcriptional gene silencing by double-stranded RNA. *Nat. Rev. Genet. 2*, 110-119.

Hutvagner, G., McLachlan, J. et al. (2001). A cellular function for the RNA-interference enzyme Dicer in the maturation of the let-7 small temporal RNA. *Science 293*, 834-838.

Ketting, R. F., Fischer, S. E. et al. (2001). Dicer functions in RNA interference and in synthesis of small RNA involved in developmental timing in C. elegans. *Genes Dev. 15*, 2654-2659.

Lagos-Quintana, M., Rauhut, R. et al. (2001). Identification of novel genes coding for small expressed RNAs. *Science 294*, 853-858.

Lau, N. C., Lim, L. P. et al. (2001). An abundant class of tiny RNAs with probable regulatory roles in Caenorhabditis elegans. *Science 294*, 858-862.

Lee, R. C. and Ambros V. (2001). An extensive class of small RNAs in Caenorhabditis elegans. *Science 294*, 862-864.

Lee, Y., Jeon, K. et al. (2002). MicroRNA maturation: stepwise processing and subcellular localization. *Embo. J. 21*, 4663-4670.

Price, L. S., Hajdo-Milasinovic, A. et al. (2004). Rap1 regulates E-cadherin-mediated cell-cell adhesion. *J Biol. Chem. 279(34)*, 35127-35132.

Reinhart, B. J., Weinstein, E. G. et al. (2002). MicroRNAs in plants. *Genes Dev. 16*, 1616-1626.

Slattery C., Campbell, E., McMorrow, T. and Ryan, M. P. (2005). Cyclosporine a-induced renal fibrosis: a role for epithelial-mesenchymal transition. *Am. J. Pathol. 167*, 395-407.

Tuschl, T., Zamore, P. D. et al. (1999). Targeted mRNA degradation by double-stranded RNA in vitro. *Genes Dev. 13*, 3191-3197.

Zamore, P. D., Tuschl, T. et al. (2000). RNAi: double-stranded RNA directs the ATP-dependent cleavage of mRNA at 21 to 23 nucleotide intervals. *Cell 101*, 25-33.

Acknowledgments

This work is supported by the Conway Institute and the Dublin Molecular Medicine Centre (DMMC) under the Programme for Research in Third Level Institutions (PRTLI) administered by the Higher Education Authority (HEA). This work is also supported by Enterprise Ireland, the Health Research Board of Ireland, Science Foundation Ireland and a research bursary from the Irish Nephrological Society and Amgen Ltd.

Correspondence to

Dr. Tara McMorrow
Department of Pharmacology
Conway Institute
School of Biomolecular and Biomedical Research
UCD Dublin, Belfield, Dublin 4
Ireland
e mail: tara.mcmorrow@ucd.ie

Detecting Neurotoxicity Through Electrical Activity Changes of Neuronal Networks on Multielectrode Neurochips*

Alexandra Gramowski[1], Simone Stüwe[1], Konstantin Jügelt[1], Dietmar Schiffmann[1], Jan Loock[2], Olaf Schröder[3], Guenter W. Gross[4] and Dieter G. Weiss[1]

[1]Institute of Biological Sciences, Department of Cell Biology and Biosystems Technology, [2]Clinic for Internal Medicine, University of Rostock, Germany, [3]Pattern Expert, Borsdorf, Germany, [4]Center for Network Neuroscience, University of North Texas, Denton, TX, USA

Summary

The innovative alternative method of multielectrode neurochip recording, which offers improvements compared to standard animal experimentation, has been further optimised by advanced multi-parametric data analysis and used to study toxicity patterns of known toxicants. Neuronal networks on 64-electrode neurochips develop long-term stable spontaneous electrical activity. Changes in spatio-temporal electrical activity were quantified and subjected to pattern analysis to evaluate the effects of selected neurotoxic compounds. Several examples underline the suitability of this broadband biosensor system for functional neurotoxicity testing and safety pharmacology. Our results demonstrate that neuronal networks retain tissue-specificity and respond to transmitter receptor blockers and other neurotoxic compounds in a substance-specific, dose-dependent manner.

Keywords: functional neurotoxicity, multielectrode arrays, multiparametric data analysis, primary neuronal networks

Abbreviations: TMT, trimethyltin chloride; CV, coefficient of variation; GABA, γ-aminobutyric acid

Introduction

Spontaneously active networks in culture have been proposed as a sensitive and efficient model system to study the neurotoxic properties of chemicals, as biosensors, and to accelerate drug development (Nelson, 1978; Gross, 1994; Gross et al., 1997). In contrast to patch-clamp and imaging methods, the neurochip technology allows online and real-time analysis of up to 256 neurons in short- and long-term studies.

Primary cultures of central nervous tissue on multielectrode neurochips offer the potential to study neuroactive properties of compounds that are technically difficult or impossible to obtain *in vivo*. *In vitro* studies can replace animal experiments that would often deliver only qualitative data. Neuronal networks represent the functional units of information processing in the brain. Due to their high level of complexity they reflect major aspects of neuronal function in mammals. Multielectrode neurochip recordings allow detection of functional deviations in developing, adult and ageing brain regions and provide insights into the mechanisms of action of neuroactive compounds.

Here we investigated whether primary neuronal networks from frontal cortex and spinal cord retain the specific receptor endowments of their different parent tissues by applying toxins acting as antagonists at the γ-aminobutyric acid (GABA) and glycine receptors. The antifouling agent trimethyltin chloride (TMT) was chosen as a representative neurotoxicant in animal models (Wenger et al., 1982; Chang and Dyer, 1983; Aschner and Aschner, 1992) to demonstrate that our method is suitable to determine the neurotoxic potential of exogenous substances (Gramowski et al., 2000). Additionally, the neurotoxic potential of human blood preparations from hepato-encephalopathic coma patients containing complex mixtures of endogenous neurotoxic metabolites (Loock et al., 2004) was studied in order to detect their influence on network electrical activity.

Materials and methods

Chemicals and test samples

The chemicals bicuculline methiodide, 5-fluoro-2'-deoxyuridine + uridine (FDU), picrotoxin, poly-D-lysine, strychnine hydrochloride and trimethyltin chloride were from Sigma-Aldrich Chemical GmbH (Taufkirchen, Germany). DNase I (from bovine pancreas), laminin and papain were purchased from Roche (Mannheim, Germany), foetal bovine serum from Pan Biotech GmbH (Aidenbach, Germany). Horse serum, Minimum Essential Medium (MEM) and Dulbecco's Modified Essential Medium (DMEM) were from GIBCO BRL (Paisley, UK). Blood plasma samples from hepato-encephalopathic patients were obtained during an FDA approval study and were provided by Teraklin AG (Rostock, Germany).

* Awarded with "The Best Poster Award" at the 5th World Congress on Alternatives and Animal Use in the Life Sciences, Berlin 2005

Multielectrode neurochips

Multielectrode neurochips were provided by the Center for Network Neuroscience (CNNS) at the University of North Texas. The fabrication techniques and the culture methods have been described previously (Gross et al., 1985; Gross, 1994). Briefly, commercially available, sputtered indium tin oxide (ITO) plates were photoetched, cut into 5x5 cm^2 plates, and spin-insulated with polysiloxane. The electrode tips were de-insulated and electroplated with a thin layer of gold to lower the interface impedance. Flaming through a stainless steel mask created a confined adhesive region (1-4 mm diameter islands centred on the electrode array), which was coated with poly-D-lysine (25 µg/ml; 30-70kD) and laminin (16 µg/ml).

Cell culture

Cortical and spinal cord tissues were harvested from embryonic day 16 and 14 NMRI mice (Charles River Inc., Sulzfeld, Germany), respectively. The mice were sacrificed by cervical dislocation in compliance with the German Animal Protection Act. The tissue was cultured according to the method of Ransom et al. (1977) with minor modifications that included the use of DNase I (8000 units/ml) and papain (10 U/ml) for tissue dissociation. The tissue was dissociated enzymatically with papain and mechanically with transfer pipettes, and seeded in DMEM (Auditory and Frontal Cortex) or MEM (Spinal Cord) 10/10 (10% horse and 10% foetal calf serum) at a density of $0.5x10^6$ cells/cm^2 onto neurochips (Gross et al., 1993; Gramowski et al., 2004). Cultures were incubated at 37°C in a 10% CO_2 atmosphere until ready for use, usually three weeks to three months after seeding (fig. 1). Culture media were replenished three times a week with DMEM or MEM containing 10% horse serum. The networks develop from a mixture of different types of postmitotic neurons and glia cells. The glia cells have important auxiliary functions regarding metabolism and supply the neurons with ions and nutrients. The co-culture of neurons with glia cells ensures stability over many months. Within the first week (usually after 3–4 days *in vitro*), the developing cultures were treated with 5-fluoro-2'-deoxyuridine (25 µM) and uridine (63 µM) for 48 h to prevent further glia proliferation. Primary spinal cord, frontal and auditory cortex networks cultured on multielectrode neurochips routinely develop spontaneous electrical activity. Activity starts after approximately three to four days *in vitro* in the form of random spiking which, after 3 weeks in culture, stabilises into complex activity patterns, composed of coordinated bursting and interburst spiking (Gramowski et al., 2004). Such networks can remain spontaneously active and pharmacologically responsive for more than six months (Gross, 1994). For this study, cultures between 26 and 113 days *in vitro* were used.

Microscopy

Networks were studied by DIC Nomarski, conventional and confocal scanning immunofluorescence microscopy and diaminobenzidine (DAB) staining for neuronal or glial markers with an Eclipse 800 from Nikon (Düsseldorf, Germany) and TCS-SP2 from Leica (Bensheim, Germany) at the Rostock University Live Cell Imaging Center. Scanning electron microscopy was performed with a Zeiss DSM 960A (Oberkochen, Germany) at the Rostock University Electron-Microscopy Center (EMZ).

Extracellular recording

The multielectrode neurochips were maintained in a recording chamber (Gross, 1994) at 37°C under humidified 10% CO_2 atmosphere to ensure a constant pH of 7.4. Recording was performed with the Multichannel Acquisition Processor System, a computer-controlled 64-channel amplifier system from Plexon Inc. (Dallas, TX, USA) providing programmable amplification, filtering, switching, and digital signal processing of multielectrode signals. The total system gain used was 10 K with a simultaneous 40 kHz sampling rate. The multichannel signal acquisition system delivered single neuron spike data. Spike identification and separation were accomplished with a template-matching algorithm in real time (fig. 2). This was also used to determine action potential shape and amplitude changes during the course of an experiment.

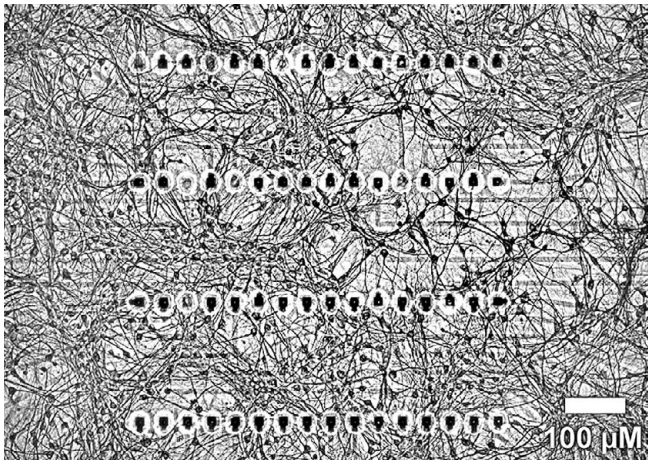


Fig. 1: Dissociated frontal cortex network on a multielectrode neurochip after 21 days *in vitro*. The culture was stained with DAB against neurofilament and MAP2.

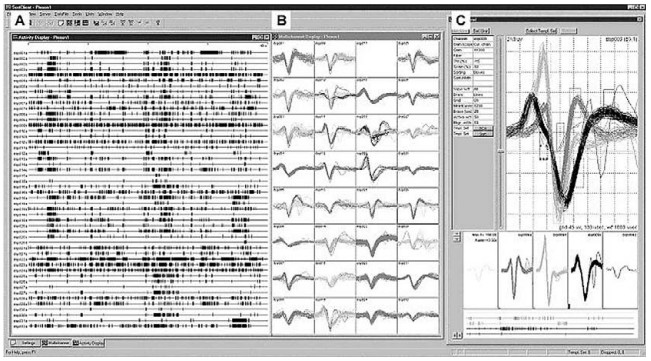


Fig. 2: Graphical user interface displaying online the native electrical activity of a spinal cord culture. (A) 40 seconds of neuronal activity. Each action potential is represented by a vertical tick. (B) Overlaid action potentials at active electrodes. (C) Separation of multiple neurons (bottom) from one electrode (top), which captures the signals from several neurons.

Data analysis

Spikes in spike trains are clustered in so-called bursts. These represent the ubiquitous and prominent feature of network activity and are an essential aspect of the neuronal code (Gross, 1994; Lisman, 1997). We extracted a total of 67 activity describing parameters to quantify the substance specific activity changes: 31 parameters derived from spike train data, their derived 31 normalised values to compensate for the variability of the activity in different cultures and five parameters derived from the dose-response curves. Bursts were determined by spike train analysis based on spike intervals with the software NeuroEXplorer (Plexon Inc., Dallas, TX, USA). Spike rate and burst rate values were derived from 60 s bin data. The dose response curves were fitted to the Hill equation, determining the effective concentration causing 50% of the maximal response (EC_{50}). The spike and burst parameters were additionally processed to derive temporal and network coefficients of variation (CV_{TIME} and $CV_{NETWORK}$) for various activity features. These CVs were used to describe the spatiotemporal behaviour of the network activity (Keefer et al., 2001; Gramowski et al., 2004). CV_{TIME} reflects the periodic behaviour of a single neuron's activity pattern and was already used in earlier studies, analysing single channel recording activity (Bracci et al., 1996). $CV_{NETWORK}$ reveals the coordination between different neurons in a specific activity state and is a measure of synchronicity.

Statistics

Results are expressed as series means ± SEM. The features' distributions were tested for normality. The level of significance after compound application was assessed using Student's paired t-test. For assessing the difference between sets of data, Student's unpaired t-test was used. $P < 0.05$ was considered statistically significant.

Results

In vivo like tissue specificity

We compared the spiking and bursting activity of spinal cord and frontal cortex networks to demonstrate that *in vitro* networks are pharmacologically histiotypic representations of their parent tissues. Neurotoxins with known site of action, namely the inhibitory $GABA_A$ and glycin receptor systems were used. Disinhibition of the electrical activity was achieved by blocking the $GABA_A$ receptors with 40 µM bicuculline ("BCC" activity) or the glycine receptors with 1 µM strychnine ("STR" activity), while complete disinhibition was reached with both compounds ("BCC+STR" activity).

Blocking the inhibitory circuits with bicuculline and/or strychnine elicited rhythmic and synchronised activity with clear-cut differences between both tissues. Strychnine increased the burst rate of spinal cord networks to 815% of native activity, while in frontal cortex networks it increased only to 127%. Bicuculline induced an increase in spike rate to 457% for spinal cord networks, compared to 204% in the case of frontal cortex networks (fig. 3). Thus, inhibition in frontal cortex cultures is mainly mediated by $GABA_A$ receptors, but in spinal cord cultures by glycine receptors. This reflects the *in vivo* situation and demonstrates that important aspects of tissue specificity are retained *in vitro*. The 7-fold increase in spike rate in spinal cord networks after complete blockade of inhibitory receptors indicates that spinal cord is more mediated by inhibitory circuits than frontal cortex with a twofold increase.

Blocking specific receptors

To demonstrate the high sensitivity of the network activity changes to specific receptor blockers we chose the $GABA_A$ receptor antagonists bicuculline and picrotoxin. Both inhibit

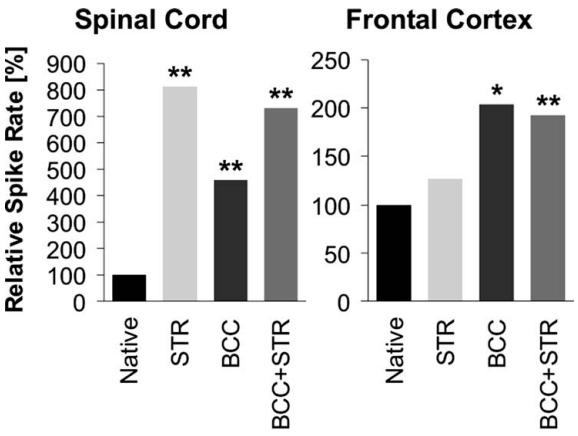


Fig. 3: Tissue-specific responses of spinal cord and frontal cortex networks. Spiking activity increases after blockade of the inhibitory $GABA_A$ receptors with 40 µM bicuculline (BCC) and/or glycine receptors with 1µM strychnine (STR). There is a stronger influence of GABA in the frontal cortex in contrast to a stronger influence of glycine on the spinal cord. (*, $p<0.05$; **, $p<0.01$).

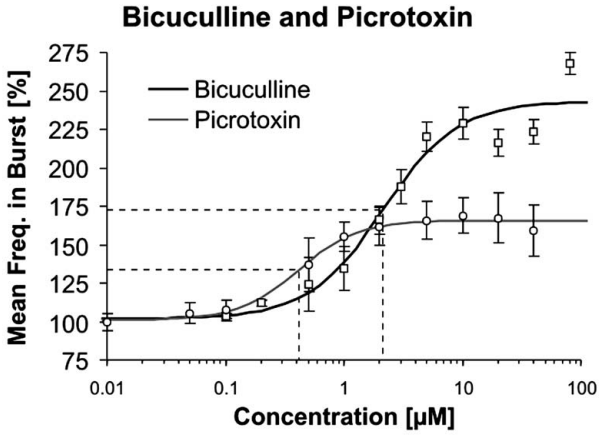


Fig. 4: Dose response curves for the two $GABA_A$ receptor antagonists bicuculline (n=6) and picrotoxin (n=7). Although in both cases $GABA_A$ receptor mediated inhibition is blocked, the different modes of action cause different responses. The changes of the mean frequency of spikes in bursts distinguish the two neurotoxins.

$GABA_A$ receptor transmission. Bicuculline blocks the GABA binding site, whereas picrotoxin clogs the Cl^- channel. Both compounds caused an increase in overall network activity with EC_{50} values of 2.1 µM for bicuculline and 0.4 µM for picrotoxin (fig. 4). They also increased network activity synchronisation and oscillation and enhanced bursting by increasing spike frequency in bursts. However, bicuculline made the burst shape significantly more uniform than picrotoxin. This was revealed by a lower CV_{TIME} of spikes in bursts. Since these severe deviations from native activity patterns were caused by the neurotoxins, such disturbances are indicators of functional neurotoxicity.

Tissue-specific neurotoxicity

Two different tissues of the mouse CNS (spinal cord and auditory cortex) exhibited characteristic and dose-dependent changes of their electrophysiological activity patterns after treatment with trimethyltin chloride (TMT), a standard neurotoxicant. Spinal cord networks began to respond at 1-2 µM, and shut off activity at 4-7 µM TMT. Auditory cortex cultures first responded at 2-3 µM; shut-off occurred at 7-8 µM TMT. Repeated applications of low doses of TMT always influenced the electrical activity in a reversible manner, with no overt cytotoxic effects. The EC_{50} values were 1.5 ± 0.5 µM (spinal cord) and 4.3 ± 0.9 µM (auditory cortex), indicating a relatively low variability among different networks of one tissue type (fig. 5). The non-overlapping EC_{50} range for cortical and spinal cord cultures suggests tissue specificity for network responses to TMT. Shut-off concentrations are within a factor of two of the lethal concentrations reported for mice *in vivo*. Action potential amplitude and shape did not change even when complete cessation of activity was approached, suggesting that acute TMT applications did not affect neuronal metabolism that may lead to a lowering of membrane potentials (fig. 6).

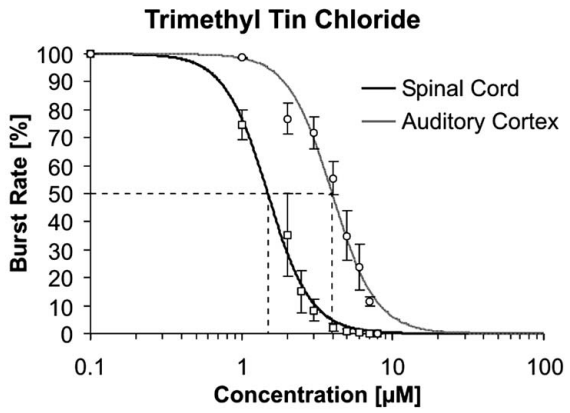


Fig. 5: Dose response curves of trimethyltin chloride on 5 spinal cord and 3 auditory cortex networks. EC_{50} values are 1.5 ± 0.5 µM for spinal cord and 4.3 ± 0.9 µM for auditory cortex, demonstrating tissue-specific sensitivity to this neurotoxicant.

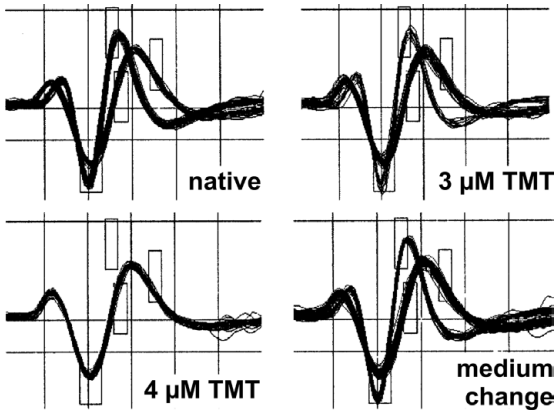


Fig. 6: Multiple superimposed action potential traces from two spinal cord neurons. They were simultaneously recorded from one electrode showing action potential stability despite exposure to 3 and 4 µM trimethyltin chloride (TMT) as well as reversal of neuronal shut-off after adding fresh medium (lower panels). These signals measured 405 and 309 µV peak-to-peak with a duration of 1.5 ms.

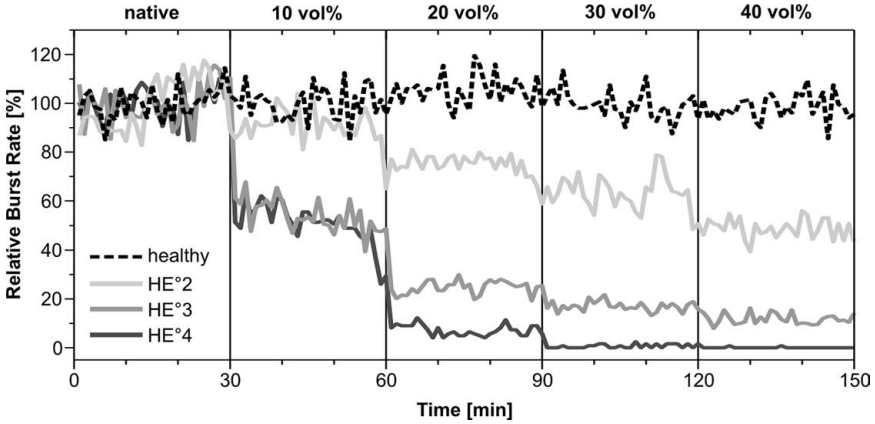


Fig. 7: Burst rate from four networks with application of blood ultrafiltrate from healthy volunteers and patients with grades of hepatic encephalopathy from HE grade 2 to 4. The burst rate decreases with increasing hepatic coma stage and volume percentage of ultrafiltrate in the culture medium.

Endogenous neurotoxins

Our system was used to monitor the CNS intoxication in hepatic encephalopathy patients. In liver failure, water-soluble metabolites as well as predominantly protein-bound compounds accumulate in the blood plasma and promote the development of encephalopathy and coma. Plasma ultrafiltrates contain the water-soluble compounds, while plasma extracts contain the protein-bound compounds. Therefore, both blood preparations were tested for their influence on network activity due to neurotoxic compounds. Plasma ultrafiltrates and extracts from healthy subjects caused only minor changes to the network activity. In contrast, ultrafiltrates from coma patients with hepatic encephalopathy (stage II-IV) predominantly decreased spike rate and burst rate (fig. 7) in a dose-dependent manner. Extracts of plasma from such patients altered network activity in a similar manner and additionally increased burst duration (data not shown).

Discussion and conclusion

Nervous tissues express patterns of electrical activity as part of their normal function. Any major interference with these patterns can generate behavioural and/or autonomic malfunctions in response to a toxic agent. The cessation of electrical activity, even if not associated with cell death ("functional neurotoxicity"), can lead to the death of the organism. Also, major changes in pattern generation can severely alter the performance of organisms without necessarily threatening survival. This is generally categorised as "behavioural toxicity" (Fiedler et al., 1996). Our studies primarily deal with functional neurotoxicity during acute exposure. The detected similarity in concentration ranges for the *in vivo* and *in vitro* situation suggests that neuronal networks *in vitro* allow a quantitative assessment of functional neurotoxicity. Whole network responses are closer to the animal situation than data from single cell studies.

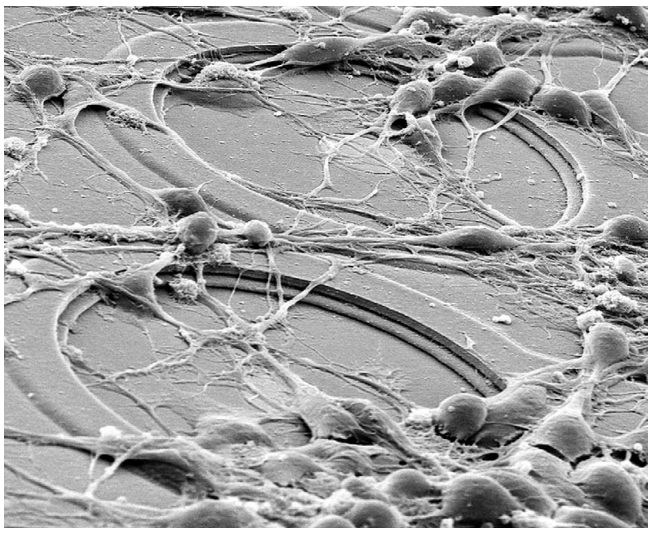


Fig. 8: SEM picture of a neuronal network growing on palladium electrodes of a CMOS silicon neurochip two days after seeding.

The multielectrode neurochip recording from functional neuronal networks is the experimental approach of choice when modifications of activity patterns by compounds are to be detected. A refined approach that we use for data analysis is performed with methods of pattern recognition, which allows the analysis of the significance and relevance of the multitude of computed features (up to 200) that are used to quantitatively describe the activity patterns. In this way a better understanding of features that are significant to describe toxicity is possible by using feature score and feature selection methods well known from pattern recognition. In this manner evidence of several burst features such as shut-off, changes in bursting patterns, in rhythmicity, overexcitation, EC_{50} values or reversibility are assigned to known toxins to characterise significant changes. In a second step the toxicity of substances is classified by their spike train responses and used to judge the toxic potential of unknown substances.

A quantitative database of neuroactive effects in specific nerve tissues *in vitro* provides a tool to classify known and unknown neuroactive compounds. Detailed knowledge of the *in vitro* effects will greatly contribute to risk assessment, although it is difficult to predict which specific symptoms in whole animals may be correlated with the different features of network activity. However, we have previously discussed how this system can be used to generate databases of well-characterised substance "fingerprints", which allow detailed comparisons of the activity spectra of neurotoxic substances (Gramowski et al., 2004).

A new generation of multi-sensor neurochips in silicon CMOS (Complementary Metal Oxide Semiconductor) technology will provide a further considerable improvement of this approach (fig. 8). It allows, in addition, online monitoring of the metabolic state of the cell system by integrated physiological sensors for temperature, oxygen consumption and pH value (Baumann et al., 2004; Krause et al., 2006).

3Rs

The multielectrode neurochip technology offers a *refinement* by giving multiparametric, fingerprint-like descriptions of the action of toxic compounds on nervous tissues. In addition, it reduces the number of animals required for neurotoxicity testing. From one parent mouse, neuronal tissue for 40-50 neurochips can be derived. Further, with each of these neurochips, dose response series with 10 to 15 concentration steps, single dose repeats, or chronic experiments can be carried out. This will reduce the number of animal experiments needed for the study of toxic compounds by a factor of up to 500 and at the same time yield much more detailed information on the mode of action of the compounds. Additionally, cultures from knock-out mice can *replace in vivo* animal disease models.

Different from patch clamp and brain slice techniques, this approach allows long-term studies (for weeks and more) as well as the testing of complex mixtures and unknown compounds. It is closer to the real situation in the nervous system, because it makes accessible the complexity level of multicellular functional ensembles, and not only that of single cells or single channels. This is also a reason why this technique is expected to give fewer false negative results than other *in vitro* techniques.

Other applications, especially in pharmacology and drug development, are presently also studied and they will, in addition to toxicity testing, open additional avenues of improved drug research and animal experiment reduction (Gramowski et al., 2004; Stüwe et al., 2005). Therefore, the application of the multielectrode neurochip technology will certainly contribute to the 3Rs of alternative methods for animal experiments and at the same time improve the quality of risk assessment and consumer protection.

References

Aschner, M. and Aschner, J. L. (1992). Cellular and molecular effects of trimethyltin and triethyltin: relevance to organotin neurotoxicity. *Neurosci. Biobehav. Rev. 16*, 427-435.

Baumann, W. H., Schreiber, E., Krause, G. et al. (2004). Cell monitoring system with multiparametric CMOS sensorchips, Proceedings µ-TAS, Malmö, Sept. 2004, Vol. II, 554-556.

Bracci, E., Ballerini, L. and Nistri A. (1996). Spontaneous rhythmic bursts induced by pharmacological block of inhibition in lumbar motoneurons of the neonatal rat spinal cord. *J. Neurophysiol. 75*, 640-647.

Chang, L. W. and Dyer, R. (1983). A time-course study of trimethyltin induced neuropathology in rats. *Neurobehav. Toxicol. Teratol. 5*, 346-350.

Fiedler, N., Feldman, R. G., Jacobson, J. et al. (1996) The assessment of neurobehavioral toxicity.: SGOMSEC joint report. *Environmental Health Perspect. 104, Suppl. 2*, 179-191.

Gramowski, A., Jügelt, K., Weiss, D. G. and Gross, G. W. (2004). Substance identification by quantitative characterization of oscillatory activity in murine spinal cord networks on microelectrode arrays. *Eur. J. Neurosci. 19*, 2815-2825.

Gramowski, A., Schiffmann, D. and Gross, G. W. (2000). Quantification of acute neurotoxic effects of trimethyltin using neuronal networks cultured on microelectrode arrays. *NeuroToxicology 21*, 331-342.

Gross, G. W. (1994). Internal dynamics of randomized mammalian neuronal networks in culture. In: D.A. Stenger and T.M. McKenna (eds.), *Enabling Technologies for Cultured Neural Networks* (277-317). New York: Academic Press.

Gross, G. W., Norton, S., Gopal, K. et al. (1997) Nerve cell network in vitro: Applications to neurotoxicology, drug development, and biosensors. *Cellular Engineering 2*, 138-147.

Gross, G. W., Rhoades, B. K., Reust, D. L. and Schwalm, F. U. (1993). Stimulation of monolayer networks in culture through thin-film indium-tin oxide recording electrodes. *J. Neurosci. Methods 50(2)*, 131-43.

Gross, G. W., Wen, W. and Lin, J. (1985). Transparent indium-tin oxide patterns for extracellular, multisite recording in neuronal cultures. *J. Neurosci. Meth. 15*, 243-252.

Keefer, E. W., Gramowski, A. and Gross, G. W. (2001). NMDA receptor-dependent periodic oscillations in cultured spinal cord networks. *J. Neurophysiol. 86*, 3030-3042.

Krause, G., Lehmann, S., Lehmann, M. et al. (2006). Measurement of electrical activity of long-term mammalian neuronal networks on semiconductor neurosensor chips and comparison with conventional microelectrode arrays. *Biosensors & Bioelectronics 21(7)*, 1272-1282.

Lisman, J. E. (1997). Bursts as a unit of neural information: making unreliable synapses reliable. *Trends Neurosci. 20*, 38-43.

Loock, J., Gramowski, A., Börner, K. et al. (2004). Assessment of the toxic load of blood in severe liver failure by responses of frontal cortex network activity to blood serum filtrates in vitro. *J. Hepatol. 40, Suppl. 1*, 70.

Nelson, P. G. (1978). Neuronal cell cultures as toxicologic test system. *Environmental Health Perspectives 26*, 125-133.

Ransom, B. R., Neale, E., Henkart, M. et al. (1977). Mouse spinal cord in cell culture. I. Morphologie and intrinsic neuronal electrophysiologic properties. *J. Neurophysiol. 40*, 1132-1150.

Stüwe, S., Gramowski, A., Schröder, O. et al. (2005). Evaluation of neuroactive compounds with neuronal networks in vitro on microelectrode neuro-sensor chips. *ALTEX 22, Special Issue*, 308.

Wenger, G. R., McMillan, D. E. and Chang, L. W. (1982). Behavioral toxicology of acute trimethyltin exposure in the mouse. *Neurobehav. Toxicol. Teratol. 4*, 157-161.

Acknowledgements

We thank Kristine Gürtler for technical assistance and Dr. L. Jonas and his collaborators for help with the electron microscopy. This work was supported by the DFG Innovationskolleg "Komplexe und Zelluläre Sensorsysteme" (INK 27), by the DARPA Tissue-Based Biosensor Program, and by the State Ministries of Education and of Economy of Mecklenburg-Vorpommern and the European Community (ERDF) through the State Priority Research Programme "Biosystems Technology".

Correspondence to

Dr. Alexandra Gramowski
University of Rostock
Dept. of Cell Biology and Biosystems Technology
Albert-Einstein-Str.3
18051 Rostock
Germany
e-mail: alexandra.gramowski@uni-rostock.de

Session 7.3
The contribution of the "OMICS"-technology to the 3Rs

Detection of Pain and Stress by Monitoring Gene Expression

Paolo Cinelli, Margarete Arras and Kurt Bürki
University of Zurich, Institute of Laboratory Animal Science, Zurich, Switzerland

Summary
One of the main concerns with regard to animal experimentation is that if animals must be used for experimental purposes, pain and distress should be abolished or reduced to an absolute minimum. Regarding the principles of the 3Rs, considerable progress has been made concerning reduction and replacement, however to implement refinement it is necessary to improve our ability to objectively recognise signs of pain and distress.
Here we present preliminary results indicating that gene expression analysis with DNA microarrays results in characteristic molecular signatures, which allow the identification of different pain and stress levels in mice.

Keywords: pain, stress, microarray, gene expression

Introduction and questions

The clear definition and monitoring of pain and stress is an essential aspect in the refinement of animal experimentation, but it is still one of the most contradictory topics in science. Regarding the principles of the 3Rs, considerable progress has been made concerning reduction and replacement, however to implement refinement it is necessary to improve our ability to objectively recognise signs of pain and distress. Detection of pain/stress in laboratory mice is very important, not only from the welfare point of view but also because pain and fear, together with the resulting stress, may introduce confounding variability to scientific data. However, the objective definition of pain/stress in animals is very difficult. The measurement of physiological parameters like corticosteroids, glucose, growth hormone or prolactin in plasma implicates handling of the animals by the researcher, a process that directly induces supplementary stress to the animals (Seggie and Brown, 1975; Tabata et al., 1998; Zethof et al., 1994). On the other hand, the telemetry technology allows measurements of physiological parameters like heart rate, body temperature and blood pressure for long periods of time without direct interference. This system allows the detection of changes in the circadian rhythm or in the activity of the animals. This system was for example used to show that mice housed individually have a higher heart rate compared with mice housed in pairs (Spani et al., 2003). Nevertheless, this is a very expensive technology and the implantation of the transmitter is bound to painful surgical procedures that induce pain for several days and require analgesic treatment of the animals.

From the behavioural point of view it is difficult to detect distress in the mouse. The mouse lives in constant fear of falling prey to its enemies, therefore it shows as few signs of disease, suffering or weakness as possible. Accordingly, during animal experiments, or even when an investigator is simply present in the room, the mouse will hide almost all signs of light or middle grade pain. In consequence, there are no reliable indicators to detect low and middle grade pain in the mouse. Only when an animal is almost moribund can we recognise that it is suffering pain. As a result, demands for adequate pain therapy are often

ignored or met with the anthropomorphic attitude “as long as the mouse shows no pain, it must be feeling no pain”.

Pain consists of both a sensory (discriminative) and an affective (the "unpleasantness") dimension, and can be acute or chronic depending on its duration. Chronic pain states can lead to secondary effects such as anxiety, depression and stress (Price, 2000).

Nociceptive information reaches the brain from the peripheral site of injury through multiple neuronal pathways (ascending pathways). Cells in the terminal regions within the forebrain and brainstem almost invariably project back, directly or indirectly, to the areas of origin of the ascending pathways. Areas like hypothalamus and amygdala are thought to play an important role in the modulation of the affective dimensions of pain and the control of the autonomic activity (Lovick, 1996; Lumb and Lovick, 1993).

Nociception is a complicated process consisting of short and long-term responses that serve as a warning, activated in response to impending damage to the organism. The threshold for eliciting pain has to be high enough so that it does not interfere with normal activities, but low enough to prevent tissue damage. This threshold is not fixed and can be shifted up or down, and it is this plasticity of the sensory system that essentially characterises pain syndromes (Woolf and Salter, 2000). Early short-term responses following nociceptor activation are reflected in rapid changes of neuronal discharge activity in a variety of anatomically distinct systems in the central nervous system. These are followed by long-term changes that most commonly require alterations in gene expression (for a review see (Scholz and Woolf, 2002)). The activity-dependent modulation of gene expression is a characteristic feature of highly integrated systems such as pain. Different molecules are involved at different levels of this process. Neurotransmitters, peptide hormones, neurosteroids, trophic factors or/and cytokines can be released from neurons, glial cells or components of the immune system and are involved in the integration of somatosensory information. Posttranslational and transcriptional changes can drastically change the threshold, excitability and transmission properties of the nociceptors. Especially in long-term changes, activity-dependent signal transduction cascades and signalling pathways downstream of the receptors bound by cytokines and growth factors act to modify transcription in nociceptor neurons. This changes the activities of the neurons, especially their transduction, conduction and transmission properties.

A major goal of pain research at the present time is the identification of pain genes. In the past few years, molecular genetic techniques allowed the study of pain at the level of the gene. Through different techniques, like the generation and analysis of transgenic mice, antisense knockdowns, DNA microarray-based expression profiling, and linkage mapping, it was possible to identify different genes directly involved in the processing of pain. Many different strains of knockout and overexpressing mice have been evaluated for their nociceptive sensitivity, both directly as the focus of the study and indirectly as part of standard behavioural screening protocols. The findings from many of these studies have been reviewed (Mogil and Grisel, 1998; Mogil et al., 2000).

These genes are interesting targets for anti-pain treatment, but are also of outstanding interest as diagnostic markers to monitor pain. Monitoring many genes at the same time became possible with the development of DNA microarray technology. DNA microarrays are among the most powerful and versatile tools of genomics and genetics research (Fodor et al., 1993; Lockhart et al., 1996; Lockhart and Winzeler, 2000; Southern et al., 1994). DNA arrays allow us to make quantitative parallel measurements of gene expression (mRNA abundance) for tens of thousands of genes. There are two dominant types of arrays that have been used for gene expression measurements. The first are high-density oligonucleotide arrays directly synthesised on a glass surface using light-directed combinatorial synthesis (Fodor et al., 1991) containing thousands of oligonucleotide sequences. The other main array type is produced by spotting cDNAs, PCR products or oligonucleotides at specific locations on a glass slide and is called low-density array.

A diagnostic microarray that can be used to monitor pain in the mouse and is specific for stress and pain does not exist yet. Only attempts to evaluate differential gene expression profiles in animal models of pain have been published previously, but no microarrays for the direct characterisation of pain itself have been described to date (Ko et al., 2002; Saban et al., 2002; Sun et al., 2002; Wang et al., 2002). Here we present the possibility to use DNA microarrays as a rapid, reliable and objective method to assess pain and stress in mice on the molecular level.

Results

Design of a low-density microarray for the detection of pain in mice

As a first step in the design of a low-density microarray, we started a careful search of the available literature for all genes related to pain, stress, and anxiety. We collected about 250 genes that directly or indirectly correlate with pain/stress. Of these 250 genes, 130 were selected for spotting on a low-density microarray. We decided to spot 70-nucleotide-long oligomers on the microarrays instead of the classical PCR-amplified gene fragments. The 70mer technology allows minimisation of the secondary structure, high melting temperatures, and therefore a normalised hybridisation temperature. We designed the 70mers by choosing their sequences within the last 750 nucleotides of the genes (3’-end), all oligonucleotides were checked by blast analysis to confirm that they are representative for the specified gene.

Definition of pain models and time points for the validation of the microarrays

In order to validate the low-density microarray, we decided to use clearly defined time points in which animals experience pain. For this we decided to use a telemetry-based postoperative pain model. Monitoring of the surgical pain by telemetry allows us not only to determine whether the animals are feeling pain, but also to determine the time window in which gene expression changes can be analysed. The pain model used consists of two different experiments. First, a moderate/strong pain is induced

by the implantation of the transmitter. During this time, heart rate, body temperature, and activity of the mouse showed strong impairment of the circadian rhythms in the animals. A tendency to reach normal values is normally detected after 1-2 weeks. After 6 weeks of recovery, the mouse was subjected to a vasectomy (mild pain). A significant increase in heart rate could be measured one day after the operation. Tissues for microarray analysis were collected five days after the transmitter implantation (moderate pain) and one day after vasectomy (mild pain). At these time points different tissues (spinal cord, brain) were collected and total RNA was purified for the analysis on microarrays.

Labelling and hybridisation of the microarrays

Brains and spinal cords were homogenised and total RNA was isolated. Labelling was performed by incorporation of aminoallyl-UTP (aa-UTP) in antisense RNA (aRNA) and subsequent labelling with cy-hydroxysuccinimide ester. This protocol allows amplification of small amounts of sample targets and was sufficiently sensitive for the analysis of brain and spinal cord from mice after vasectomy and transmitter implantation. After comparison of the data obtained from probes after transmitter implantation and after vasectomy, we could identify about 60 genes that displayed clearly modulated expression in mice with pain (tab.1).

Discussion

All experimental work with animals has to be monitored by careful assessment and minimisation of pain and stress. The same holds true for breeding of mutant animals. Pain and stress conditions are not only linked to the experimental procedures, but often also depend on the genetic mutations that the animals carry. However, especially in genetically modified animals, it is very difficult to correctly identify distress and pain. Often it is impossible to predict whether the animals will feel pain or will be stressed, because of insufficient knowledge on the gene that is being modified.

The assessment of pain and stress in laboratory rodents is difficult and often subject to painful investigations (e.g. telemetry studies). Telemetry is a very reliable technology, but requires experienced personnel for the implantation of the transmitters, a long time for the analysis of the measurements and is very expensive. For these reasons it is not possible to use telemetry as a routine diagnostic method for the detection of pain.

Microarray analysis in contrast is a very powerful technique that allows the fast and reliable measurement of gene expression changes for hundreds of genes in a relatively short time.

Microarray analysis can be used alone or in parallel with behavioural observations in order to clearly define stress. By using post-mortem biopsy material, the microarray technology is – from the point of view of animal welfare – much gentler, because it does not require any surgical implantation into the animals and the animals thus do not need to recover from surgery.

The data presented here represent a preliminary analysis based on a single pain model (postoperative pain model). Further experiments with tissues from other pain models will be necessary to finally validate this diagnostic tool. In general these data suggest that it is possible to identify specific molecular signatures and that these can be considered a valuable tool for effective pain detection and control.

References

Fodor, S. P., Rava, R. P., Huang, X. C. et al. (1993). Multiplexed biochemical assays with biological chips. *Nature 364*, 555-556.

Fodor, S. P., Read, J. L., Pirrung, M. C. et al. (1991). Light-directed, spatially addressable parallel chemical synthesis. *Science 251*, 767-773.

Ko, J., Na, D. S., Lee, Y. H. et al. (2002). cDNA microarray analysis of the differential gene expression in the neuropathic pain and electroacupuncture treatment models. *J. Biochem. Mol. Biol. 35*, 420-427.

Lockhart, D. J., Dong, H., Byrne, M. C. et al. (1996). Expression monitoring by hybridization to high-density oligonucleotide arrays. *Nat. Biotechnol. 14*, 1675-1680.

Lockhart, D. J. and Winzeler, E. A. (2000). Genomics, gene expression and DNA arrays. *Nature 405*, 827-836.

Lovick, T. A. (1996). Midbrain and medullary regulation of defensive cardiovascular functions. *Prog. Brain Res. 107*, 301-313.

Lumb, B. M. and Lovick, T. A. (1993). The rostral hypothalamus: an area for the integration of autonomic and sensory responsiveness. *J. Neurophysiol. 70*, 1570-1577.

Tab. 1: Genes spotted on the low-density microarray and genes differentially expressed in a model for surgical pain. Genes were considered as representative for pain when the expression ratio between control animals and animals with pain was more than two-fold. Not listed are control genes needed for normalisation and for checking the quality of the hybridisation.

Gene family	No. of genes present on the array	Genes involved in surgical pain
Receptors	51	31
Neuropeptides	10	6
Oxidative stress	9	4
Hormone-related genes	22	11
Neurotrophins	4	2
Heat shock	5	3
Channels	2	1
Cytokine-related genes	8	3

Mogil, J. S. and Grisel, J. E. (1998). Transgenic studies of pain. *Pain 77*, 107-128.

Mogil, J. S., Yu, L. and Basbaum, A. I. (2000). Pain genes?: natural variation and transgenic mutants. *Annu. Rev. Neurosci. 23*, 777-811.

Price, D. D. (2000). Psychological and neural mechanisms of the affective dimension of pain. *Science 288*, 1769-1772.

Saban, M. R., Nguyen, N. B., Hammond, T. G. and Saban, R. (2002). Gene expression profiling of mouse bladder inflammatory responses to LPS, substance P, and antigen-stimulation. *Am. J. Pathol. 160*, 2095-2110.

Scholz, J. and Woolf, C. J. (2002). Can we conquer pain? *Nat. Neurosci. 5 Suppl*, 1062-1067.

Seggie, J. A. and Brown, G. M. (1975). Stress response patterns of plasma corticosterone, prolactin, and growth hormone in the rat, following handling or exposure to novel environment. *Can. J. Physiol. Pharmacol. 53*, 629-637.

Southern, E. M., Case-Green, S. C., Elder, J. K. et al. (1994). Arrays of complementary oligonucleotides for analysing the hybridisation behaviour of nucleic acids. *Nucleic. Acids. Res. 22*, 1368-1373.

Spani, D., Arras, M., Konig, B. and Rulicke, T. (2003). Higher heart rate of laboratory mice housed individually vs in pairs. *Lab. Anim. 37*, 54-62.

Sun, H., Xu, J., Della Penna, K. B. et al. (2002). Dorsal horn-enriched genes identified by DNA microarray, in situ hybridization and immunohistochemistry. *BMC Neurosci. 3*, 11.

Tabata, H., Kitamura, T. and Nagamatsu, N. (1998). Comparison of effects of restraint, cage transportation, anaesthesia and repeated bleeding on plasma glucose levels between mice and rats. *Lab. Anim. 32*, 143-148.

Wang, H., Sun, H., Della Penna, K. et al. (2002). Chronic neuropathic pain is accompanied by global changes in gene expression and shares pathobiology with neurodegenerative diseases. *Neuroscience 114*, 529-546.

Woolf, C. J. and Salter, M. W. (2000). Neuronal plasticity: increasing the gain in pain. *Science 288*, 1765-1769.

Zethof, T. J., Van der Heyden, J. A., Tolboom, J. T. and Olivier, B. (1994). Stress-induced hyperthermia in mice: a methodological study. *Physiol. Behav. 55*, 109-115.

Acknowledgements

This work was supported by the "Forschungskredit" of the University of Zurich and by the "Foundation Research 3R", Switzerland.

Correspondence to

Dr. Paolo Cinelli
Institute of Laboratory Animal Science
University of Zurich
Winterthurerstrasse 190
8057 Zurich
Switzerland
paolo.cinelli@ltk.unizh.ch

Session 7.4
Non-invasive techniques for monitoring and imaging (Doerenkamp-Zbinden Symposium)

Magnetic Resonance Imaging (MRI) of the Lung as a Tool for the Non-Invasive Evaluation of Drugs in Rat Models of Airways Diseases

Nicolau Beckmann[1], Harry Karmouty Quintana[1,2], François-Xavier Blé[1,3], Catherine Cannet[1], Stefan Zurbruegg[1], Bruno Tigani[1] and John Fozard[1]

[1]Novartis Institutes for BioMedical Research, Discovery Technologies and Respiratory Diseases Department, Basel, Switzerland; [2]King's College London, Sackler Institute of Pulmonary Pharmacology, UK; [3]Université Louis Pasteur-Strasbourg-1, Faculté de Pharmacie, Illkirch Cedex, France

Summary

Methods currently used to evaluate the efficacy of potential treatments for diseases of the airways in small animal models are generally invasive and terminal. We explored the flexibility of magnetic resonance imaging (MRI) to obtain anatomical and functional information on the lung, with the scope of developing a non-invasive approach for the routine testing of drugs in rat models of airways diseases. With MRI, the disease progression can be followed in the same animal. Thus, a significant reduction in the number of animals used for experimentation may be achieved, as well as minimal interference with their well-being and physiological status. Also, MRI has the potential to shorten the overall duration of the observation period after disease onset, since the technique is able to detect changes induced by allergen before these are reflected in secreted parameters of inflammation.

Keywords: asthma, chronic obstructive pulmonary disease, COPD, drug development, lung, magnetic resonance imaging, MRI

Introduction

Diseases of the airways such as asthma and chronic obstructive pulmonary disease (COPD) involve a complex interplay of many inflammatory and structural cell types, all of which can release inflammatory mediators including cytokines, chemokines, growth factors, and adhesion molecules. Activated eosinophils are considered particularly important in asthma, contributing to epithelial cell damage, bronchial hyperresponsiveness, plasma exudation and oedema of the airway mucosa, as well as smooth muscle hypertrophy and mucus plugging, through the release of enzymes and proteins (Kroegel et al., 1994; Barnes, 1996; Moqbel, 1996). In COPD, inflammation of the small airways and lung parenchyma with the involvement of neutrophils, macrophages and cytotoxic (CD8+) T lymphocytes results in chronic obstructive bronchitis, destruction of the lung parenchyma by proteolytic enzymes (emphysema), and mucus hypersecretion, leading to severe airflow limitation (Barnes, 2002; Sethi, 2005).

Many species of animals can be used to provide models for airways diseases in humans. Unfortunately, there is no animal model that exactly reproduces the human pathology. Nonetheless, these models are useful for the development of novel therapies and to mimic and study specific aspects of human respiratory diseases (Dawkins and Stockley, 2001; Canning, 2003; Isenberg-Feig et al., 2003). Actively sensitised Brown Norway (BN) rats exposed to allergen develop airway hyperresponsiveness and eosinophilic inflammation together with an increase in activated T cells (CD25+) in the airways (Renzi et al., 1993; Haczku et al., 1997;

Hannon et al., 2001), hence reflecting the key features of asthmatic inflammation. Alternatively, inflammation similar to that observed in COPD patients can be elicited in rodents by the administration of endotoxin (lipopolysaccharide; LPS), a bacterial macromolecular cell wall component. LPS activates mononuclear phagocytes through a receptor-mediated process, leading to the release of a number of cytokines, including tumour necrosis factor-α (TNF-α) (Watson et al., 1994; Yang et al., 1998). TNF-α increases the adherence of neutrophils to endothelial cells, therefore facilitating a massive infiltration of neutrophils into the lung (Albelda et al., 1994). Exposure of BN rats to LPS leads to pulmonary neutrophilia (Tesfaigzi et al., 2000; Tigani et al., 2002) and induces mucus cell metaplasia (Harkema and Hotchkiss, 1992).

Invasive and terminal approaches (broncho-alveolar lavage (BAL) fluid analysis; bronchial biopsies; histology; weighing of lungs) are currently used to analyse such models and to perform preclinical drug studies. Evidently, for ethical reasons, it would be highly desirable to have non-invasive readouts in this research area.

Although *in vivo* MR techniques have been in use in pharmaceutical research for more than 20 years (for reviews, see Rudin et al., 1999; Beckmann et al., 2001a, 2004a), they have only recently been applied to pre-clinical studies in the area of respiratory diseases (Beckmann et al., 2003). Interestingly, a similar time-lag is evident for the clinical application of lung MRI. Probably the main reasons for this delay are the inherent difficulties imposed by the lung tissue on the MR signal properties and, in the clinical arena, the fact that computerised tomography is the imaging technique of choice for diagnosis of lung diseases. The aim of this chapter is to illustrate how the flexibility of MRI can be exploited to non-invasively derive information on lung inflammation and on its functional status in models of airways diseases in rats, and how this information can be ultimately used to profile compounds in these animal models (Beckmann et al., 2003).

Animals, materials and methods

Procedures are described in detail in (Beckmann et al., 2001b, 2002; Tigani et al., 2002). BN rats weighing 250-300 g were used. Two models of pulmonary inflammation were studied by MRI:

Inflammation models: (i) Allergen-induced pulmonary inflammation in actively sensitised rats, resulting from the intra-tracheal (i.t.) administration of OVA. (ii) Endotoxin-induced pulmonary inflammation in non-sensitised naïve rats, caused by instillation of LPS.

MRI: Measurements were carried out with a spectrometer operating at 4.7 T. For image acquisition, rats were anaesthetised with 2% isoflurane in a mixture of O_2/N_2O, administered via a face mask. All measurements were performed in spontaneously breathing animals; neither cardiac nor respiratory triggering was applied.

Analysis of BAL fluid or histology: Carried out in order to better understand the signals detected by MRI in the lungs. These

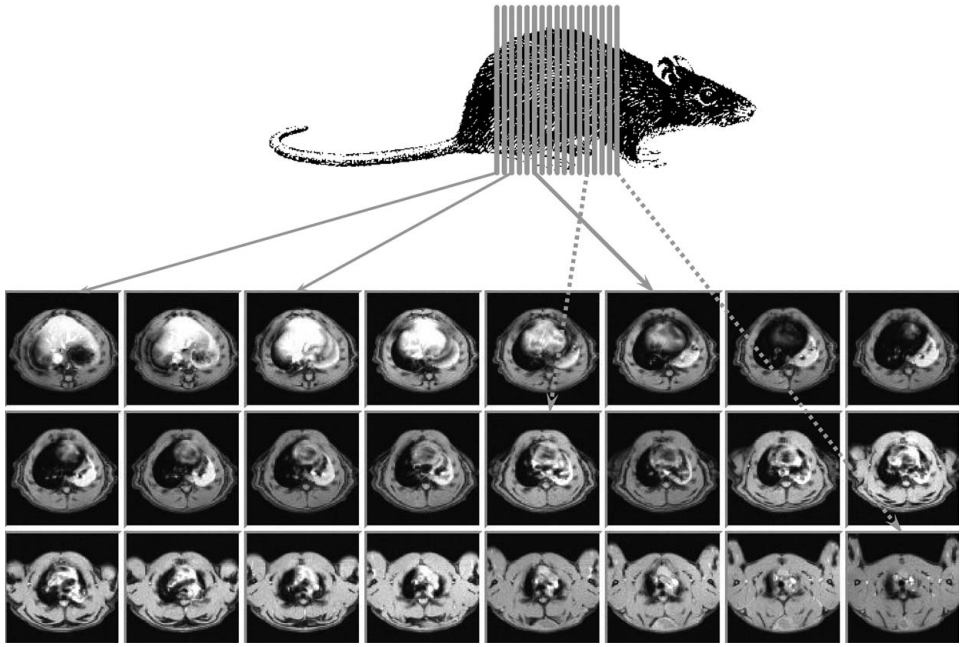


Fig. 1: Axial MR images acquired sequentially through the chest of an actively sensitised BN rat, 24 h after i.t. OVA (0.3 mg/kg) challenge, using a gradient-echo sequence. The acquisition time per image was 60 s. The animal was kept under anaesthesia (isoflurane administered through a mask) and respired spontaneously. Neither respiratory nor cardiac gating was applied during image acquisition. Oedema area is assessed on each image using a semi-automatic segmentation procedure. Total oedema volume is computed by adding the areas obtained for each slice, multiplied by the slice thickness. See Beckmann et al. (2001b) for more details.

terminal methods were only used during the establishment phase of MRI within lung research. The lungs were lavaged and several inflammation parameters were assessed in the BAL, such as cellular infiltration, including eosinophil and neutrophil numbers. Histology was carried out to assess, e.g. perivascular oedema, mucus and goblet cell numbers.

Results

Lung inflammation

A characteristic feature of respiratory diseases such as asthma is oedema of the airways due to an increase in the permeability of the lung microvasculature. The resulting effect is the leakage of fluid containing plasma proteins from the microvascular circulation into the surrounding tissue. Assessment of this fluid can be important for diagnostic purposes and for planning and guiding treatment.

Proton MRI is the natural candidate for trying to detect inflammatory responses in the lungs in models of airways diseases. We used a conventional gradient-echo technique to generate images of the rat thorax in which motion artefacts were suppressed by averaging (Beckmann et al., 2001b). Neither respiratory nor cardiac triggering was applied, and the animals were able to breathe spontaneously during data collection. Under the conditions chosen for acquisition, the signal from the lung parenchyma itself is too weak to be detected at 4.7 T. However, the absence of any detectable lung parenchymal signal in combination with a background devoid of artefacts provided a high contrast-to-noise ratio for the detection of fluid signals associated with the inflammatory process (Beckmann et al., 2001b, 2002).

In rats actively sensitised to ovalbumin (OVA) and challenged with the antigen, an intense, uniform oedematous signal was detected in the lungs 24 h after challenge. By acquiring approximately 20 images displaced from each other by a distance corresponding to a single slice thickness, the whole thorax of the animal was scanned and the total volume of the oedematous signal determined (fig. 1). The volume of the oedematous signal was dependent on the dose of allergen and reached a maximum 48 h after challenge (fig. 2). The MRI signal could be seen for approximately 100 h and was correlated highly significantly with a variety of inflammatory parameters determined in the BAL fluid recovered from the same animals (Tigani et al., 2002).

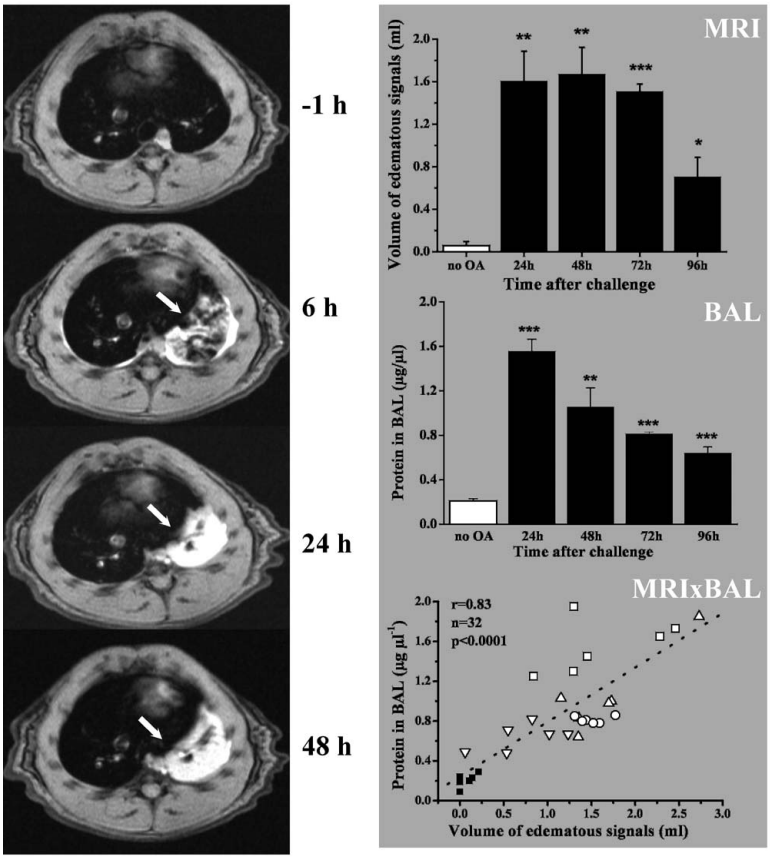


Fig. 2: Axial MR images acquired from an actively sensitised BN rat at different time points with respect to i.t. OVA (0.3 mg/kg) challenge. Prominent oedematous signals were seen already 6 h after challenge. The curves on the right depict the time course of MRI signals (mean ± SEM, n=6), and for the same animals, the protein (mean ± SEM) assessed in BAL fluid. The MRI signals significantly correlated with the protein and several markers of inflammation in the BAL. For more details, see Beckmann et al. (2001b) and Tigani et al. (2002).

Of special interest, the strongest correlations were with the eosinophil numbers, eosinophil peroxidase activity (a marker of eosinophil activation), and the total protein concentration (a marker for plasma extravasation). Importantly, the signal detected by MRI correlated significantly with the perivascular oedema assessed by histology (Tigani et al., 2003a).

Following challenge of non-sensitised rats with LPS, the signals that appeared in the lungs were uneven and significantly less intense than those detected after OVA administration to actively sensitised animals (fig. 3). They were of long duration, being detectable up to 8 days after dosing (Beckmann et al., 2002). The only parameter in the BAL fluid that correlated significantly with the MRI signal was the mucus concentration (Beckmann et al., 2002; Tigani et al., 2002). Histological analysis indicated a substantial and sustained increase in goblet cell numbers up to 16 days after LPS challenge, and flocculent mucoid material was consistently detected close to the apical surface of epithelial cells (Beckmann et al., 2002). These observations suggest that the long lasting MRI signal following LPS was due to secreted mucus.

Lung ventilation

Gas exchange is the major function of the lungs. The regional pulmonary blood flow (perfusion) and the ventilation in the lungs need to be matched for this process to occur efficiently. In diseases of the airways like asthma and COPD, lung ventilation can be compromised. Asthma is a highly prevalent chronic inflammatory disorder of the airways characterised by periodic and reversible narrowing of the airways making breathing difficult. Bronchial hyperresponsiveness to various stimuli such as irritants, infection, exercise, cold air, or allergens is a key feature of asthma and is related to an enhanced sensitivity of the airway smooth muscle to contractile stimuli. In COPD, sustained smoking causes chronic inflammation of the airways responsible for mucosal thickening, airway narrowing and loss of elastic recoil.

Pulmonary function tests providing global lung function information, such as forced expired volume in one second (FEV1) or flow-volume pattern, are used to quantify the severity of lung diseases and to evaluate the efficiency of treatment. Invasive measurements of airflow and transpulmonary pressure following stimulus by agents inducing bronchoconstriction are used to monitor lung function in rats (Hannon et al., 2001). However, these tests do not provide any regional information. The following subsections address efforts made to obtain regional information about lung function in rat models non-invasively by using proton MRI.

Airway remodelling and hyporesponsiveness induced by inflammation

Inflammation of the airways leads to pathophysiological changes in the structure of the lung tissue, including thickening of the airway smooth muscle (Ebina et al., 1993), which may influence the responsiveness to bronchospasmogens (Martin et al., 2000) and alter ventilation. The progressive structural change known as airway remodelling, which is driven by chronic local inflammation, is a fundamental component in the development of airway hyperresponsiveness (for a recent review, see Halayko and Amrani, 2003).

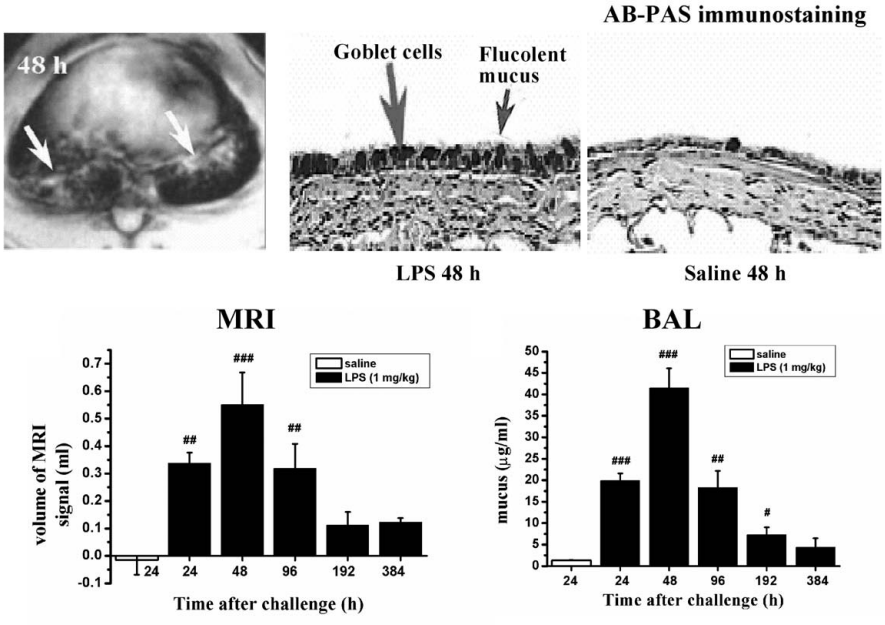


Fig. 3: Axial image through the chest of a naïve BN rat, 48 h after i.t. instillation of LPS (1 mg/kg). Histology and BAL fluid analysis revealed that the MR signal reflected secreted mucus. For details, see Beckmann et al. (2002) and Tigani et al. (2002).

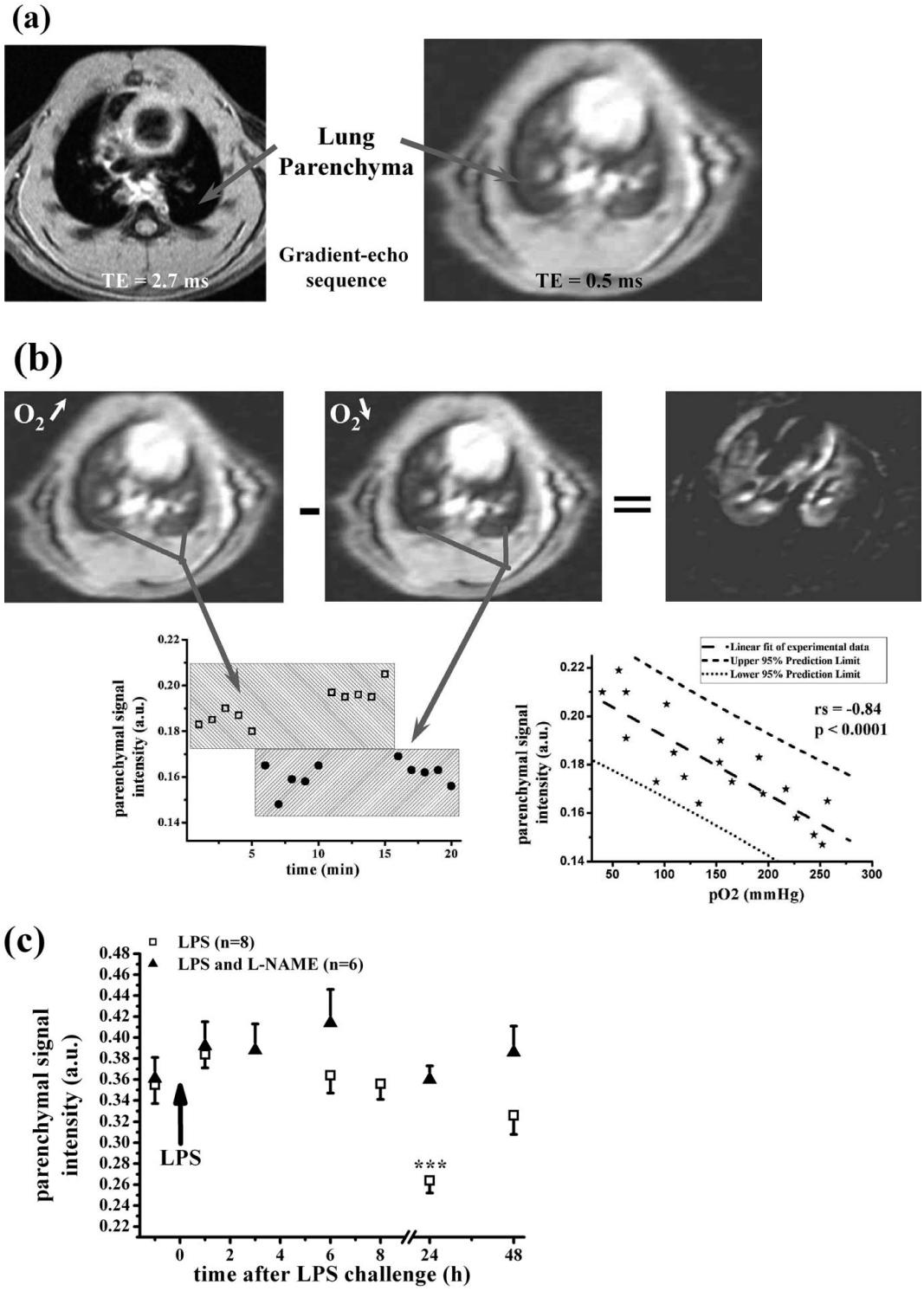


Fig. 4: (a) By changing the acquisition parameters of the gradient-echo sequence, signals from lung parenchyma become more evident. The rat could respire spontaneously, and neither cardiac nor respiratory gating was applied during image acquisition. (b) The parenchymal signal intensity is inversely dependent on the amount of O_2 in the lung. The reason is that molecular O_2 is paramagnetic, thus acting as a natural contrast agent. We found a significant negative correlation between the parenchymal signal intensity and the partial oxygen pressure in blood for different amounts of oxygen administered (between 21% and 65%). (c) Course of lung parenchymal signal (mean ± SEM) following challenge with LPS (1 mg/kg i.t.). Some of the animals were pre-treated with L-NAME (10 mg/kg i.v.) 10 min before the challenge. The significance level corresponds to comparisons with baseline signal intensities before challenge (*** $p<0.001$). See Beckmann et al. (2004b) for more details.

Effects of airway remodelling and hyporesponsiveness following respectively allergen or endotoxin challenges were monitored non-invasively in spontaneously breathing rats with a gradient-echo sequence as described by Beckmann et al. (2001c) (fig. 4a). The basis of the approach consisted in detecting modulations of proton signals of lung parenchyma induced by changes in oxygenation levels. An increased parenchymal signal should be consistent with a reduced oxygen level and *vice versa* (Beckmann et al., 2004b) (fig. 4b). This hypothesis has been verified in the allergen and endotoxin models of airways inflammation in the rat (Beckmann et al., 2004b). In actively sensitised rats, an increased parenchymal signal intensity (in areas devoid of oedematous signals) was detected at 6 h and up to 180 h after challenge, at a time when oedematous signals reflecting inflammation had completely subsided. Histological analysis revealed airway remodelling in the lungs of OVA-challenged rats characterised as an increased bronchial epithelium thickness and smooth muscle area, as well as bronchial goblet cell hyperplasia. Thus, the increased parenchymal signal in lung images of rats treated with allergen was consistent with a significant reduction of air space determined by histology, pointing to impaired lung ventilation in these animals. The ventilation defect was still observed after the oedematous signals detected by MRI were completely resolved (Beckmann et al., 2004b).

In a second model, significantly decreased parenchymal signal intensity was detected 24 h after intra-tracheal instillation of LPS (Beckmann et al., 2004b) (fig. 4c). The effect was abolished by pre-treatment with N^G-nitro-L-arginine methyl ester (L-NAME), an inhibitor of nitric oxide (NO) synthase (NOS). A possible role of NO in the inflammatory response elicited by endotoxin has been demonstrated by Pauwels et al. (1990), who showed that a period of significant hyporesponsiveness, characterised by reduced pulmonary resistance due to increased airway calibre, followed from 9 to 12 h after exposure of rats to aerosolised LPS. In the same model, LPS-induced airway hyporesponsiveness was eliminated by L-NAME (Kips et al., 1995). Moreover, a marked expression of inducible NOS in rat macrophages recovered from the airways has been demonstrated 16 h after local LPS instillation (Kobzik et al., 1993). Therefore, endogenous synthesis of NO, a potent bronchodilator (Abderrahmane et al., 1998), induced by endotoxin might have been responsible for increased oxygenation of the lung tissue, thus contributing to a reduction of the parenchymal signal in the gradient-echo images 24 h following LPS administration.

Airway smooth muscle contraction

Proton MRI was also used to detect the effects of bronchoconstrictor and bronchodilator compounds in spontaneously breathing rats. For instance, a significant increase of the parenchymal signal intensity was observed in the upper airways from the first minutes following *i.v.* administration of a compound eliciting bronchoconstriction. The long-lasting signal increase was reversed by application of a bronchodilator agent, consistent with an increase in oxygenation. Airway resistance measures derived invasively in anaesthetised, paralysed, and artificially ventilated rats showed the same time profile as that of the MRI signal. These observations suggest that the MRI signal changes in the upper airways were due to contraction of airway smooth muscle.

Drug treatment analysis

A variety of new compounds for the treatment of respiratory diseases are under development, many of which are designed as anti-inflammatory therapies (Barnes, 2002, 2004). Since oedema is an integral component of experimental pulmonary inflammation, MRI has the potential to provide a non-invasive means of monitoring the course of the inflammatory response and the consequence of therapy with anti-inflammatory drugs.

We consider first the classic approach of analysing the effects of anti-inflammatory drugs administered prior to allergen challenge (fig. 5a). A clear dose-related reduction of the oedematous signal has been shown for compounds such as the glucocorticosteroids budesonide (Beckmann et al., 2001b; Tigani et al., 2003a) and mometasone (Tigani et al., 2003b), and for a mitogen-activated protein kinase inhibitor (Tigani et al., 2003b). The effects correlated with changes in the parameters of inflammation assessed in the BAL fluid (Tigani et al., 2002). Repeated measurements allowed information on the duration of action to be easily defined.

The MRI technique was also applied to address the effects of drugs administered after the allergic inflammatory response had developed. In one experimental paradigm the drugs were given 24 h after OVA challenge, a time point when an extensive MRI signal was present in the rat lung (fig. 5b). Treatment with budesonide, mometasone or with a selective inhibitor of phosphodiesterase type 4 (PDE4), which is also a powerful inhibitor of allergic pulmonary inflammation in the rat (Trifilieff et al., 2002), accelerated the rate of resolution of the MRI signal (Beckmann et al., 2001b; Tigani et al., 2003a,b). For these compounds, a clear trend towards a reduction in the oedematous signal was observed as early as 3 h after drug administration, and the effect was statistically significant from 6 to 72 h. The decline in the oedematous signal correlated significantly with the reduction in perivascular oedema quantified by histology of the lungs (Tigani et al., 2003a). This suggests that suppression of perivascular oedema following "therapeutic" treatment with budesonide, mometasone or the PDE4 inhibitor caused the decrease in the MRI signal. By contrast, BAL fluid markers of inflammation were not affected by any compound 6 h after treatment (Tigani et al., 2003a). It seems, accordingly, that the early resolution of MRI oedematous signals by the anti-inflammatory drugs did not involve general suppression of the inflammatory response, at least as monitored by BAL fluid analysis. At 48 h following treatment with the steroids or with a PDE4 inhibitor, MPO activity and protein concentrations were significantly reduced and, in animals treated with the PDE4 inhibitor, eosinophil number and EPO activity were also significantly diminished (Tigani et al., 2003a). These changes may be the mechanistic basis for the sustained resolution of MRI signals.

In addition to profiling anti-inflammatory compounds, MRI can also be used to address the effects of compounds designed

to improve lung function. For instance, the examples discussed above suggest that proton MRI approaches have the potential to profile bronchodilator compounds in models of airways obstruction.

Discussion and conclusions

The main advantage of using MRI for the characterisation of animal models of diseases is its non-invasive nature, which allows repetitive measurements to be carried out in the same animal. This is of relevance in longitudinal studies, since the inter-individual variance is eliminated and the number of animals required to reach statistical significance is much smaller. Also, non-invasiveness is a major asset when studying chronic diseases. The flexibility of MRI makes it a widely applicable method. We are currently applying MRI techniques to non-invasively assess drug effects in several disease models, e.g. arthritis, transplantation, neurodegeneration, and stroke (Rudin et al., 1999; Beckmann et al., 2001a, 2004a).

In the case of rat models of asthma, we estimate a reduction by approximately 80-90% in the number of animals used in the studies as compared to BAL fluid analysis or histology. Although providing comprehensive information at the cellular level, those methods have the drawback of being terminal. Here, we have demonstrated how MRI can provide complementary information with a fundamental asset: its non-invasive character. Repeated measurements can be carried out on the same animal, and the time courses of events becomes easily accessible. The significant correlation between the MRI signals and the perivascular oedema determined histologically provides solid evidence for the non-invasive assessment of a key component of inflammation in the allergen model, enabling rapid effects of drugs to be detected *in vivo* by monitoring the rate at which oedematous signals resolve. Also, the prospect of using MRI to detect non-invasively a sustained mucus hypersecretory phenotype induced by endotoxin provides an important new perspective for animal models of COPD. Thus, despite being a macroscopic technique, MRI allows an overall assessment of the time-related behaviour of compounds in models of lung inflammation in rats. With this information it becomes easier to choose the time point for carrying out BAL fluid or histological analysis in order to obtain more specific information on the drug mechanism.

Since MRI studies are conducted on spontaneously breathing rats, the well-being of the animals during experimentation is improved, as invasive procedures like tracheotomy and/or intubation are avoided. Thus, repetitive measurements can be carried out more easily, and the information obtained from time courses provides a better picture of disease development and treatment. Furthermore, our aim is to replace currently adopted methods of ventilation assessment involving the use of radioactive materials. Finally, we estimate that by using MRI the duration of the experimental period can be reduced in some of the applications as compared to conventional approaches.

Overall, we are confident that MRI and other imaging techniques will play an increasingly important role in pre-clinical research on small rodents in the area of airways diseases. The non-invasive character of the approaches should facilitate not only drug assessments in animal models, but also provide relevant data for the transition to the clinic.

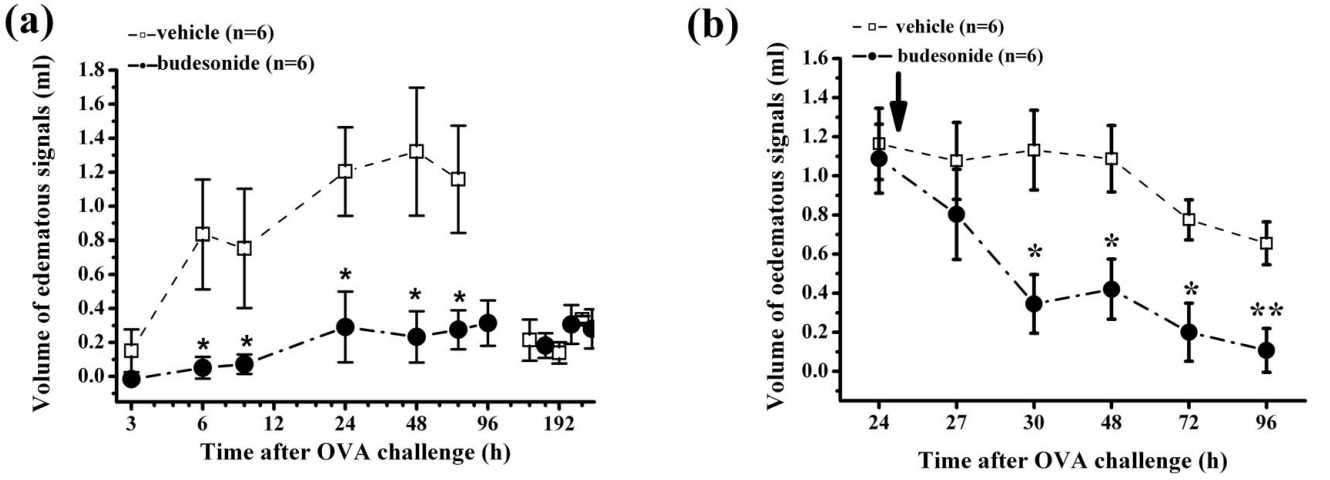


Fig. 5: Course of oedematous signals (mean ± SEM) following OVA (0.3 mg/kg i.t.) challenge in actively sensitised BN rats.
(a) Pre-treatment: animals received budesonide (1 mg/kg i.t.), a glucocorticosteroid, or its vehicle (saline) 1 h before and 24 h after OVA. (b) Post-treatment: rats received budesonide (1 mg/kg i.t.) or its vehicle 24 h after OVA. The significance levels *p<0.05 and **p<0.01 refer to t-test comparisons made between budesonide- and saline-treated animals, at each time point. For more details, see Beckmann et al. (2001b) and Tigani et al. (2003a).

References

Abderrahmane, A., Salvail, D., Dumoulin, M. et al. (1998). Direct activation of K_{Ca} channel in airway smooth muscle by nitric oxide: Involvement of a nitrothiosylation mechanism? *Am. J. Respir. Cell. Mol. Biol. 19*, 485-497.

Albelda, S. M., Smith, C. W. and Ward, P. A. (1994). Adhesion molecules and inflammatory injury. *FASEB J. 8*, 504-512.

Barnes, P. J. (1996). Pathophysiology of asthma. *Br. J. Clin. Pharmacol. 42*, 3-10.

Barnes, P. J. (2002). New Treatments for COPD. *Nat. Rev. Drug Discov. 1*, 437-446.

Barnes, P. J. (2004). New Drugs for Asthma. *Nat. Rev. Drug Discov. 3*, 831-844.

Beckmann, N., Mueggler, T., Allegrini, P. R. et al. (2001a). From anatomy to the target: contributions of magnetic resonance imaging to preclinical pharmaceutical research. *Anat. Rec. 265*, 85-100.

Beckmann, N., Tigani, B., Ekatodramis, D. et al. (2001b). Pulmonary edema induced by allergen challenge in the rat: Non-invasive assessment by magnetic resonance imaging. *Magn. Reson. Med. 45*, 88-95.

Beckmann, N., Tigani, B., Mazzoni, L. and Fozard, J. R. (2001c). MRI of lung parenchyma in rats and mice using a gradient-echo sequence. *NMR Biomed. 14*, 297-306.

Beckmann, N., Tigani, B., Sugar, R. et al. (2002). Non-invasive detection of endotoxin induced mucus hypersecretion in rat lung by magnetic resonance imaging. *Am. J. Physiol. Lung Cell. Mol. Physiol. 283*, L22-L30.

Beckmann, N., Tigani, B., Mazzoni, L. and Fozard, J. R. (2003). Magnetic resonance imaging of the lung provides potential for non-invasive preclinical evaluation of drugs. *Trends Pharmacol. Sci. 24*, 550-554.

Beckmann, N., Laurent, D., Tigani, B. et al. (2004a). Magnetic resonance imaging in drug discovery: lessons from disease areas. *Drug Discov. Today 9*, 35-42.

Beckmann, N., Cannet, C., Zurbruegg, S. et al. (2004b). Proton MRI of lung parenchyma reflects allergen-induced airway remodeling and endotoxin-aroused hyporesponsiveness: a step towards ventilation studies in spontaneously breathing rats. *Magn. Reson. Med. 52*, 258-268.

Canning, B. J. (2003). Modeling asthma and COPD in animals: a pointless exercise? *Curr. Op. Pharmacol. 3*, 244-250.

Dawkins, P. A. and Stockley, R. A. (2001). Animal models of chronic obstructive pulmonary disease. *Thorax 56*, 972-977.

Ebina, M., Takahashi, T., Chiba, T. and Motomiya, M. (1993). Cellular hypertrophy and hyperplasia of airway smooth muscles underlying bronchial asthma: a 3-D morphometric study. *Am. Rev. Respir. Dis. 148*, 720-726.

Haczku, A., Macary, P., Huang, T. J. et al. (1997). Adoptive transfer of allergen-specific CD4+ T cells induces airway inflammation and hyperresponsiveness in brown-Norway rats. *Immunology 91*, 176-185.

Halayko, A. J. and Amrani, Y. (2003). Mechanisms of inflammation-mediated airway smooth muscle plasticity and airways remodeling in asthma. *Respir. Physiol. Neurobiol. 137*, 209-222.

Hannon, J. P., Tigani, B., Williams, I. et al. (2001). Mechanism of airway hyperresponsiveness to adenoside induced by allergen challenge in actively sensitized Brown Norway rats. *Br. J. Pharmacol. 132*, 1509-1523.

Harkema, J. R. and Hotchkiss, J. A. (1992). In vivo effects of endotoxin on intraepithelial mucosubstances in rat pulmonary airways. Quantitative histochemistry. *Am. J. Pathol. 141*, 307-317.

Isenberg-Feig, H., Justice, J. P. and Keane-Myers, A. (2003). Animal models of allergic asthma. *Curr. Allergy Asthma Rep. 3*, 70-78.

Kips, J. C., Lefebvre, R. A., Peleman, R. A. et al. (1995). The effect of a nitric oxide synthase inhibitor on the modulation of airway responsiveness in rats. *Am. J. Respir. Crit. Care Med. 151*, 1165-1169.

Kobzik, L., Bredt, D. S., Lowenstein, C. J. et al. (1993). Nitric oxide synthase in human and rat lung: immunocytochemical and histochemical localization. *Am. J. Respir. Cell. Mol. Biol. 9*, 371-377.

Martin, J. G., Duguet, A. and Eidelman, D. H. (2000). The contribution of airway smooth muscle to airway narrowing and airway hyperresponsiveness in disease. *Eur. Respir. J. 16*, 349-354.

Moqbel, R. (1996). Eosinophil-derived cytokines in allergic inflammation and asthma. *Ann. N. Y. Acad. Sci. 796*, 209-217.

Kroegel, C., Virchow, J. C. Jr, Luttmann, W. et al. (1994). Pulmonary immune cells in health and disease: the eosinophil leucocyte (Part I). *Eur. Respir. J. 7*, 519-543.

Pauwels, R. A., Kips, J. C., Peleman, R. A. and van der Straeten, M. E. (1990). The effect of endotoxin inhalation on airway responsiveness and cellular influx in rats. *Am. Rev. Respir. Dis. 141*, 540-545.

Renzi, P. M., Olivenstein, R. and Martin, J. G. (1993). Inflammatory cell populations in the airways and parenchyma after antigen challenge in the rat. *Am. Rev. Resp. Dis. 147*, 967-974.

Rudin, M., Beckmann, N., Porszasz, R. et al. (1999). In vivo magnetic resonance methods in pharmaceutical research: current status and perspectives. *NMR Biomed. 12*, 69-97.

Sethi, S. (2005). Pathogenesis and treatment of acute exacerbations of chronic obstructive pulmonary disease. *Semin. Respir. Crit. Care Med. 26*, 192-203.

Tesfaigzi, Y., Fischer, M. J., Martin, A. J. and Seagrave, J. (2000). Bcl-2 in LPS- and allergen-induced hyperplastic mucous cells in airway epithelia of Brown Norway rats. *Am. J. Physiol. Lung Cell Mol. Physiol. 279*, L1210-1217.

Tigani, B., Schaeublin, E., Sugar, R. et al. (2002). Pulmonary inflammation monitored non-invasively by MRI in freely breathing rats. *Biochem. Biophys. Res. Commun. 292*, 216-221.

Tigani, B., Cannet, C., Zurbrugg, S. et al. (2003a). Resolution of the oedema associated with allergic pulmonary inflammation in rats. *Br. J. Pharmacol. 140*, 239-246.

Tigani, B., Di Padova, F., Zurbrugg, S. et al. (2003b). Effects of a mitogen-activated protein kinase inhibitor on allergic

airways inflammation in the rat studied by magnetic resonance imaging. *Eur. J. Pharmacol. 482*, 319-324.

Trifilieff, A., Wyss, D., Walker, C. et al. (2002). Pharmacological profile of a novel phosphodiesterase 4 inhibitor, 4-(8-benzo[1,2,5]oxadiazol-5-yl-[1,7]naphthyridin -6-yl)-benzoic acid (NVP-ABE171), a 1,7-naphthyridine derivative, with anti-inflammatory activities. *J. Pharmacol. Exp. Ther. 301*, 241-248.

Watson, R. W., Redmond, H. P. and Bouchier-Hayes, D. (1994). Role of endotoxin in mononuclear phagocyte-mediated inflammatory responses. *J. Leukoc. Biol. 56*, 95-103.

Yang, R. B., Mark, M. R., Gray, A. et al. (1998). Toll-like receptor-2 mediates lipopolysaccharide-induced cellular signalling. *Nature 395*, 284-288.

Acknowledgement

NB acknowledges the receipt of an award from the 3R Research Foundation, Münsingen, Switzerland (Project No. 82/02).

Abbreviations

BAL, broncho-alveolar lavage; BN, Brown Norway; COPD, chronic obstructive pulmonary disease; i.t., intra-tracheal; L-NAME, N^G-nitro-L-arginine methyl ester; LPS, lipopolysaccharide; MAP, mitogen activated protein; MR, magnetic resonance; MRI, magnetic resonance imaging; NO, nitric oxide; NOS, nitric oxide synthase; OVA, ovalbumin; PDE4, phosphodiesterase-4; TNF-α, tumor-necrosis factor-α

Correspondence to

Dr. Nicolau Beckmann
Novartis Institutes for BioMedical Research
Discovery Technologies
Analytics and Imaging Sciences Unit
Lichtstr. 35, WSJ-386.2.09
4002 Basel, Switzerland
e-mail: nicolau.beckmann@novartis.com

Session 7.5
Novel cell culture techniques

Modelling Long-Term Repeat-Dose Toxicity – Challenges Faced

Richard Clothier, Jonathan Crookdale, Jessica Morgan, Lucy Fitzgerald and Pauline Moore
FRAME Alternatives Laboratory, School of Biomedical Sciences, University of Nottingham, UK

Summary
Modelling of repeat/chronic-dose toxicity requires the in vitro maintenance of defined functions that are partly dependent on medium supplements. The ability to grow cells in the undifferentiated state and then control their differentiation is critical, since toxicity can affect differentiation triggers, differentiation control, and/or maintenance of the differentiated state.
Required monitoring endpoint assay methods should allow repeat exposure, and assess both functional and toxicological endpoints. In addition, in vitro toxicity should evaluate recovery from injury, since this is part of in vivo toxicity profiling.

Keywords: in vitro, repeat toxicity, chronic toxicity

Introduction

Modelling of repeat/chronic dose toxicity is a complex area of research, but will be restricted to consideration of epithelial barriers in this short paper. Papers from the group of Pfaller will deal with organ culture approaches (Balls et al., 2001).

Chronic and/or repeat approaches to toxicity have been examined by a number of authors (e.g., Balls et al., 1982; Dierickx and Ekwall, 1992; Clothier and Samson, 1996; Clothier et al., 1997). Problems arise not only with respect to the choice of target cells but also regarding the medium composition (Wilkinson and Clothier, 2005), since defined *in vitro* functions must be retained, which depends on certain medium supplements (Rubin and Rice, 1986) and cell stability *in vitro*.

The ability to grow epithelial cells in the undifferentiated state and then control differentiation is critical (Rheinwald and Green, 1975; Ward et al., 1997), as toxicity can affect differentiation triggers, the differentiation control process (Rubin and Rice, 1986; Gray et al., 1999), and/or maintenance of the differentiated state (Barker and Clothier, 1997).

In addition, *in vitro* toxicity should include evaluation of recovery from injury (e.g. Clothier et al., 1997, 1999a), which is part of *in vivo* toxicity profiling.

Monitoring endpoint assay methods should allow for repeat exposure of cultures, and assess both functional and toxicological endpoints. Many routinely employed and even validated assays require fixation and solubilisation of cultures, thus numerous replicate cultures must be performed (e.g. Dierickx and Ekwall, 1992; Spielmann et al., 1998). Hence, the FRAME Alternative Laboratory has been developing endpoint assays that allow repeat monitoring of the same cultures (e.g. Shaw et al., 1991; Clothier and Samson, 1996) with standardised protocols (e.g. Gray et al., 2004). New methods and imaging systems such as confocal microscopy (Garside et al., 1998) and Terahertz radiation (Arnone et al., 1999; Clothier and Bourne, 2003) are still required.

The Alamar Blue™ assay is based on the conversion of resazurin to resorufin (O'Brien et al., 2000) at concentrations that do not apparently harm the differentiation capacity of keratinocytes or corneocytes (Clothier and Samson 1996; Clothier

and Bourne, 2003; Gray et al., 2004). The enzymatic pathway of resazurin reduction is known (Andrews at al., 1997; O'Brien et al., 2000).

Non-toxic fluorescent indicator dyes, e.g. cell tracker green (CTG) that monitors glutathione (GSH), are valuable tools for the evaluation of chronic/repeat effects of chemicals on epithelial cells (Reid et al., submitted).

The importance of innervation to the repair and maintenance of differentiation, particularly in the human eye, is a focus of increasing interest (e.g. Suuronen et al., 2004; Moore et al., 2005).

Materials and methods

Human keratinocytes

were isolated and cultured according to Barker and Clothier (1997), or purchased from Cambrex UK Ltd., and cultured according to the ICCVAM/ECVAM validation study protocol (Strickland et al., 2003). Human corneal cells were cultured according to published protocols (Moore et al., 2005).

The human keratinocyte cell line (HaCaT) was cultured in keratinocyte medium (Epilife from Cascade Biologics,

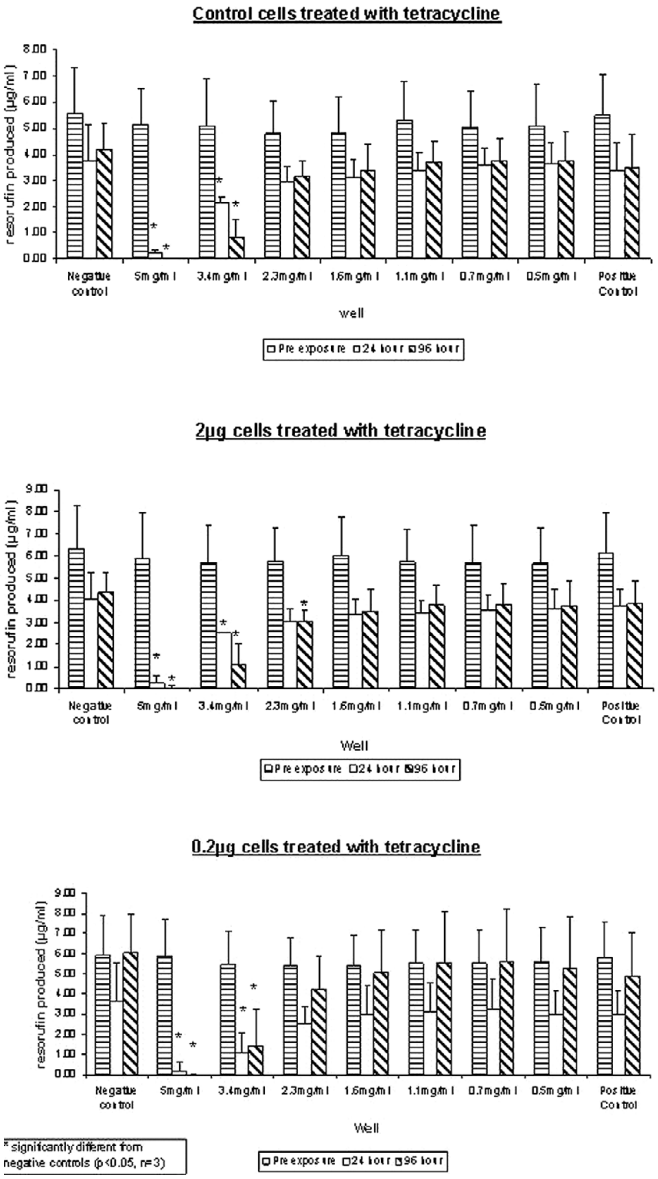


Fig. 1: Effects of pre-treatment with non-cytotoxic concentrations of tetracycline upon and acute exposure to tetracycline. Values are expressed as mean ± SD (n=3). Data analysed by repeated measures one-way ANOVA followed by Dunnett's test; * indicates significant difference from negative control, i.e non-pretreated with tetracycline.

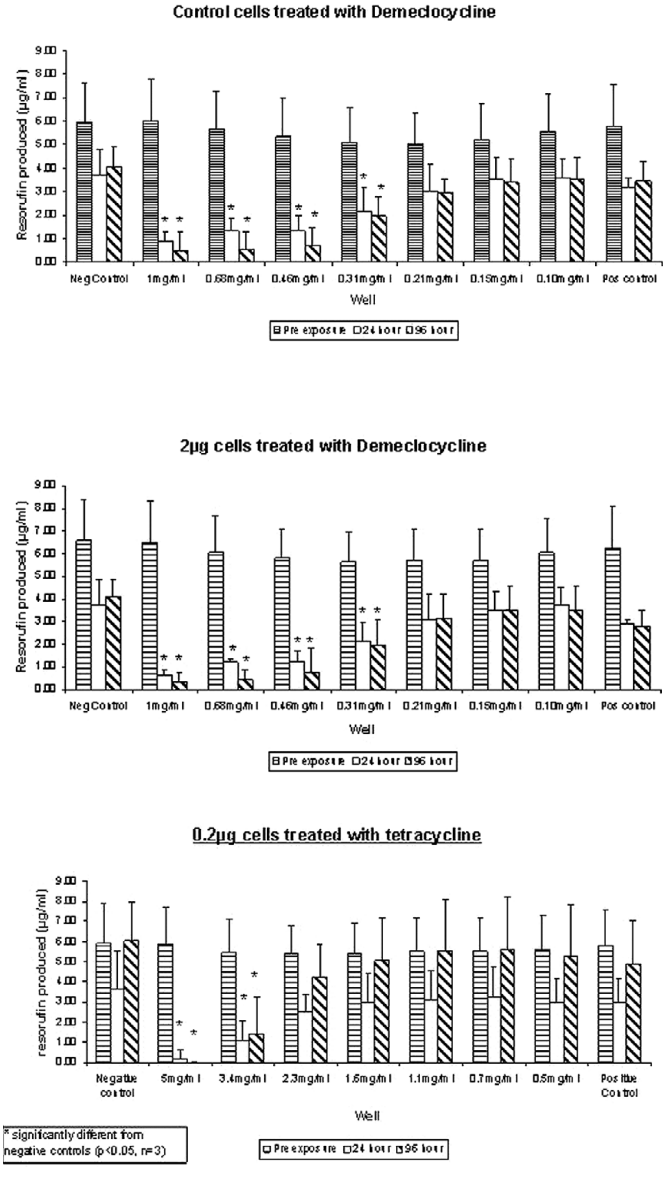


Fig. 2: Effects of pre-treatment with non-cytotoxic concentrations of tetracycline upon and acute exposure to demeclocycline. Values are expressed as mean ± SD (n=3). Data analysed by repeated measures one-way ANOVA followed by Dunnett's test; * indicates significant difference from negative control, i.e non-pretreated with demeclocycline.

Mansfield, UK). Confluent monolayers were exposed to 100 µM buthionine sulphoximide (BSO; Sigma, Poole, Dorset, UK) for 18 h. 2.2 µM bithionol (Sigma) was then added for 1 h prior to placing in the dark or exposure to UVA-visible light with a SOL-500 lamp as published previously (Clothier et al., 1999b; Reid et al., 2001).

10 µM CTG (Molecular Probes, Leiden, The Netherlands) in HBSS (Sigma, Poole, Dorset, UK) was added to cells for 20 min. Cells were photoactivated with UVA or kept in the dark and then examined at intervals using a fluorescence microscope under UV light.

Human corneal cells

(Araki Sasaki et al., 1995) were exposed to tetracycline or demeclocycline (Sigma, Poole, Dorset, UK) at 0, 0.2 or 2 µg/ml for 21 days and then transferred to the central 60 wells of a 96 well plate (Nucleon, SLS, Nottingham, UK). After 24 h, allowing for cell attachment and prior to exposure for 1 h to concentrations of 5000, 3400, 2300, 1600, 1100, 700 or 500 µg/ml tetracycline or 1000, 680, 460, 310, 210, 150, 100 µg/ml demeclocycline, cell activity was monitored using the resazurin assay (Gray et al., 2004). The assay was repeated at 24 h and 96 h post exposure on the same cells. Six replicate wells were treated with each concentration and the experiment was repeated on three separate occasions with differing passage numbers.

Co-cultures of a sensory neuronal cell line

(ND7/23; ATCC, LGC, London) with human keratinocytes, corneal cells or bronchial cells, was achieved by growing neural cells in DMEM as recommended (Life Technologies Ltd, Paisley, Scotland), and then seeding onto the underside of polycarbonate 0.45 µm tissue culture inserts (Nunc, SLS, Nottingham, UK). Cell attachment occurred within 2-3 h and cells were cultured further for up to 24 h. Medium was then changed to medium defined for epithelium cells under examination and added into the insert; method according to Moore et al. (2005).

Results

To demonstrate the effects of chronic exposure of human epithelial cells to antibiotics, corneal epithelial cells were exposed to tetracycline.

The prior exposure to low, non-cytotoxic concentrations of tetracycline affected the sensitivity of the cultures to subsequent cytotoxic concentrations of tetracycline (fig. 1). Cells pre-treated with 2 µg/ml tetracycline were sensitive to 2300 µg/ml at 96 h, whilst at this latter concentration and below those pre-treated with 0 or 0.2 µg/ml tetracycline were not affected in their vitality (fig. 1). The pre-treatment with tetracycline did not affect the cultures' sensitivity to any of the concentrations of demeclocycline (fig. 2).

Using bithionol as a toxicant and CTG as the fluorescein indicator for GSH, the ability of human keratinocytes to resist the toxic effects of photoactivated bithionol was examined.

In previous studies, when human keratinocytes were tested blindly for detection of phototoxins, bithionol was not identified as a toxin (Clothier et al., 1999b; Combes et al., 1999). However, additional "stressing" of the cells revealed the phototoxic effects (Reid et al., 2001).

Pre-exposure of keratinocytes to CTG allowed GSH location within them. Subsequent exposure to bithionol did not affect the distribution (fig. 3), unless the cells were also exposed to UVA-visible light (c.f. fig. 4 and 5). To demonstrate the toxicity of bithionol's photoproduct, cells were rendered more responsive by BSO pre-treatment which reduces GSH levels. Under such "stressed" conditions, the vesicular location of CTG was noticeable from 1.5 h, peaking at 4 h, maintained up to 24 h and lost by 29 h (fig. 6). For those cells not exposed to UVA-visible light, no such vesicular concentration was observed. The CTG remained in cells for over 96 h and could be replenished by subsequent exposure/s without adverse effects.

Changes in cell activity can also be observed in correspondingly treated cells that were not exposed to UVA but kept in the dark (Tab.1). Here, pre-exposure to the BSO resulted in reduced

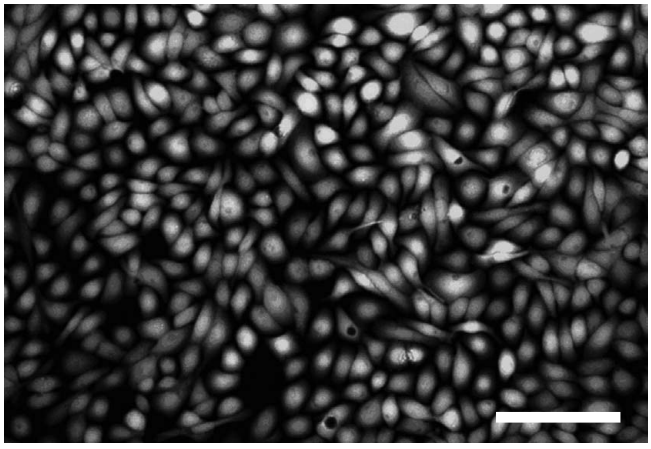


Fig. 3: Image of keratinocytes exposed to BSO followed by CTG. Bar = 30 µm

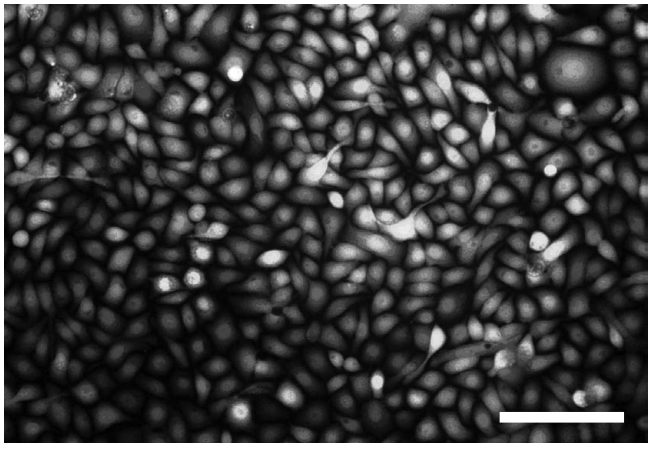


Fig. 4: Image of keratinocytes exposed to BSO followed by 2 µM Bithionol and then CTG, and kept in the dark for 24 h. Bar = 30 µm

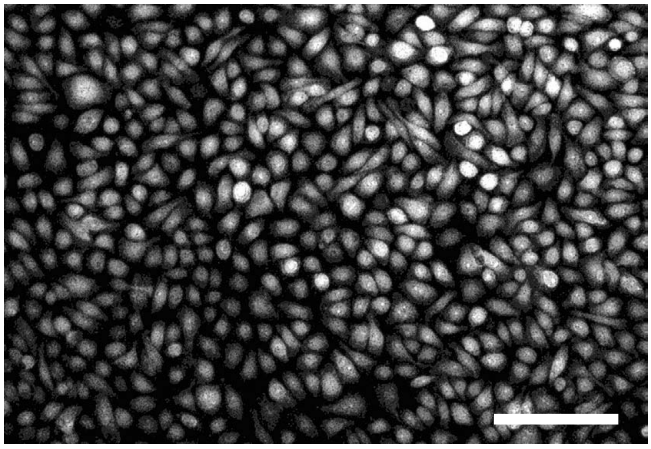


Fig. 5: Image of keratinocytes exposed to BSO followed by 2 µM Bithionol and then exposed to UVA-visible light prior to CTG. Vesiculation as seen at 4 and 24 h post UV activation. Bar = 30 µm

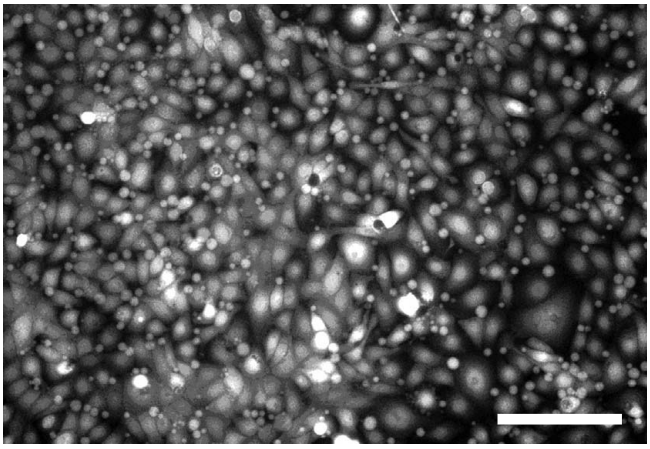


Fig. 6: Image of keratinocytes exposed to BSO followed by 2 µM Bithionol and then exposed to UVA-visible light prior to CTG. Absence of vesiculation at 29 h post UV activation. Bar = 30 µm

viability as measured using the resazurin assay. Upon exposure to 0.69 µM bithionol and UVA, a recovery of cell activity could be observed at 48 h. Hence, location of GSH, changes in cell activity and recovery could be evaluated in the same cells.

Epithelial cell recovery from injury can be modulated by the presence of sensory nerves (Moore et al., 2005), thus the ND7/23 cell line was employed to investigate its effects on barrier function. Initial resorufin production by the Japanese Human Corneal Epithelial (JHEC) cell monolayers ranged from 2.21 ± 0.60 to 2.76 ± 0.77 µg/ml/h.

Calcium concentrations in the range of 0-100 µM showed a significant effect upon cell activity 24 h after changing the medium's calcium concentration. 5 µM ($p<0.05$) calcium caused an increase in resorufin production above control level, but this was short-lived and by 48 h was similar to the control (fig. 7).

There was a concentration-dependent increase in barrier function, measured by the reduced fluorescein leakage over time, but only for calcium concentrations of 100 µM at 24 h and 50 µM and above at 72 h (fig. 7). However, when ND7/23 cells were grown with corneal cells on the underside of the inserts, sensory neurons did not compensate for the lack of calcium.

Discussion

A number of approaches with endpoint assays have been outlined that can reveal both function and structure, and allow for continuous monitoring of the same cultures pre- and post-exposure, enabling the description of recovery profiles.

Studies have demonstrated that resazurin and fluorescein leakage endpoints that examine cell activity, including NADPH-quinone oxidoreductase-dependent pathways (Andrews et al., 1997), and tight-junctional integrity (Shaw et al., 1990) can be combined and performed simultaneously (Clothier and Samson, 1996). It has also been possible to combine the neutral red uptake assay (Riddell et al., 1986) with the resazurin assay, but this unfortunately requires desorption of neutral red, terminating cell culture.

Use of confocal microscopy can also follow CTG and resazurin uptake and the temporal conversion of resazurin to resorufin, prior to and following toxicant exposure (Reid, Moore and Clothier unpublished results).

The new technologies surrounding the use of non-ionising radiation, such as the THz wavelengths (Arnone et al., 1999), also hold promise for non-invasive methods (e.g. Clothier and Bourne, 2003; Woodward et al., 2003) to detect changes in cultures following toxicant exposure, and during subsequent recovery.

Tab. 1: Effects of the reduction of GSH levels by BSO on HaCaT cell viability when treated with bithionol and then UVA at 5 J/cm²

Bithionol concentration	Non-BSO-treated 24 h	BSO-treated 24 h	Non-BSO-treated 48 h	BSO-treated 48 h
control	100 ± 30	90 ± 25	100 ± 32	97 ± 30
1.5 µM	2 ± 1	4 ± 2	3 ± 1	0
1.02 µM	30 ± 5	10 ± 5	32 ± 4	7 ± 4
0.69 µM	52 ± 10	25 ± 5	57 ± 7	37 ± 5
0.32 µM	80 ± 10	78 ± 12	85 ± 12	74 ± 15

Mean ± SD of percentage of corresponding dark controls, n=3.

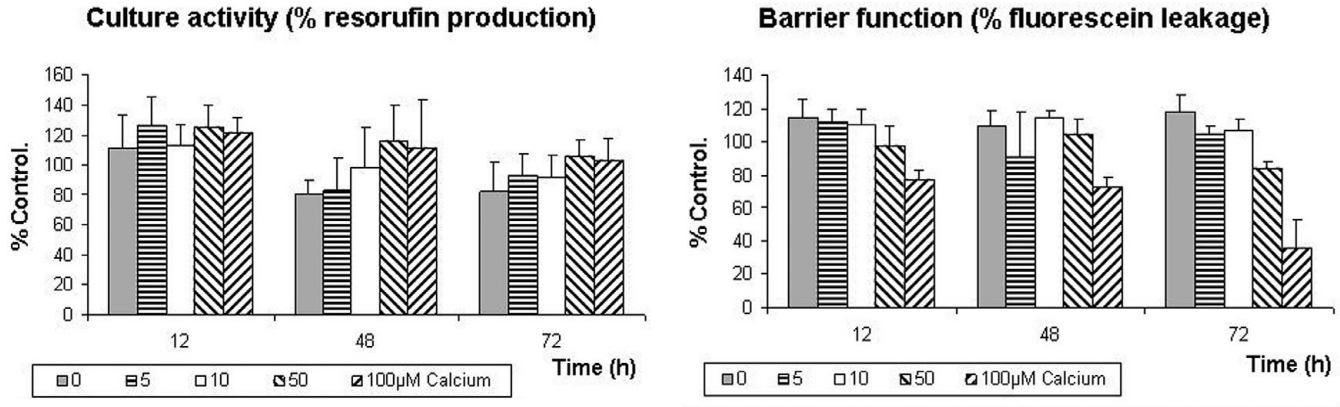


Fig. 7: The effects of 1-100 µM calcium on cell viability and barrier formation of JHCE-T cells in the presence of ND7/23 cells in the wells. Values are expressed as mean ± SD (n=3). Data analysed by repeated measures one-way ANOVA followed by Dunnett's test; ** indicates significant difference from 0 hours for the same conditions. Data analysed by repeated measures two-way ANOVA followed by Bonferroni's test; * indicates significant difference from 0 calcium at the same time point.

The co-culture approach with neural and epithelial cells is stable for many days. In addition, neural cells in the 3D epithelial models and keratinocyte medium differentiate and demonstrate the presence of ion-gated channels and neurite projection (Moore et al., 2005).

Thus this type of approach is being used to observe recovery from chronic exposure in the presence or absence of nerve cells. Consideration of such methods will assist in the development of chronic and repeat-insult models including description of recovery capacity.

References

Andrews, M. J., Garle, M. J. and Clothier, R. H. (1997). Reduction of the new tetrazolium dye Alamar blue, in cultured rat hepatocytes and liver fractions. *ATLA 25*, 641-653.

Araki-Sasaki, K. Ohashi, Y. Sasabbe, T. et al. (1995). As SV-40 imortalised human corneal epithelium cell line and its characterisation. *Investigative Ophthamology and Visual Science 36*, 614-621.

Arnone, D. D., Ciesla, C. M. Corchia, A. et al. (1999). Applications of Terahertz technology to Medical Imaging. *Terahertz Spectroscopy & Applications – II, SPIE 3828*, 209-219.

Balls, M., Santos, M. A. and Clothier, R. H. (1982). The long-term maintenance of differentiated tissues in vitro: paracetamol metabolism and toxicity in organ culture. In M. M. Yeoman and D. E. S. Truman (eds.), *Differentiation in vitro: British Society for Cell Biology Symposium 4* (25-47). Cambridge: Cambridge University Press.

Balls, M. Pfaller, W. Coecke, S. et al. (2001). Novel advanced in vitro methods for long-term toxicity testing. ECVAM Workshop report 45. *ATLA 29*, 393-426.

Barker, C. L. and Clothier, R. H. (1997). Human keratinocyte Cultures as models of cutaneous esterase activity. *Toxicology in Vitro 11*, 637-640.

Clothier, R. and Samson, R. (1996). Effects of surfactant retreatment in vitro: a method to evaluate changes in cell junctions and in cell viability. *ATLA 24*, 859-865.

Clothier, R., Beed, M., Samson, R. and Ward, R. (1997). An in vitro approach to the evaluation of repeat exposure in the prediction of toxicity. *Toxicology in Vitro 11*, 679-682.

Clothier, R., Starzec, G., Stipho, S. and Kwong, Y. C. (1999a). Assessment of initial damage and recovery following exposure of MDCK cells to an irritant. *Toxicology in Vitro 13*, 713-717.

Clothier, R. H., Willshaw, A., Cox, H. et al. (1999b). The use of human keratinocytes in the ECVAM/COLIPA in vitro phototoxicity study. *ATLA 27*, 247-259.

Clothier, R. H. and Bourne, N. (2003). Effects of THz exposure on human primary keratinocyte differentiation and viability. *Journal of Biological Physic 29*, 179-185.

Combes, R., Manners, J., Boavida, P. et al. (1999).The effects of UVA on human corneal and keratinocyte cell lines. In D, Clark, S. Lisaansky and R. Macmillian (eds), *Alternatives to Animal Testing II* (152-155). Newbury, UK: CPL Press.

Dierickx, P.J. and Ekwall, B. (1992). Long-term cytotoxicity testing of the first twenty MEIC chemicals by the determination of protein content in human embryonic lung cells. *ATLA 20*, 285-289.

Garside, J. R., Clothier, R. H., Somekh, M. G. and See, C. W. (1998).The use of a simultaneous fluorescence and differential phase confocal microscope to study Alamar Blue™ reduction in an Epithelial cell line. *ATLA 26*, 505-522.

Gray, A. C., Garle, M. and Clothier, R. H. (1999). Fluorescein cadaverine incorporation as a novel technique for the characterisation of terminal differentiation in keratinocytes. *Toxicology in Vitro 13*, 773-778.

Gray, A. C., Malton, J. and Clothier, R. H. (2004). The devel-

opment of a standardised protocol to measure squamous differentiation in stratified epithelia, using the fluorescein cadaverine incorporation technique. *ATLA 32*, 91-100.

Moore, P., Ogilvie, J., Horridge, E. et al. (2005). The development of an innervated epithelial barrier model using a human corneal cell line and ND7/23 sensory neurones. *European Journal of Cell Biology 84*, 581-592.

O'Brien, J., Wilson, I., Orton, T. and Pognan, F. (2000). Investigation of the Alamar Blue (resazurin) fluorescent dye for the assessment of mammalian cell cytotoxicity. *European Journal of Biochemistry 267*, 5421-5426.

Reid, L., Clothier, R. H. and Khammo, N. (2001). Hydrogen peroxide induced stress in human keratinocytes and its effect on bithionol toxicity. *Toxicology in Vitro 15*, 441-445.

Reid, L., Khammo, N. and Clothier, R. H. Evaluation of the effects of photoactivated bithionol in human keratinocyte in vitro. Submitted for publication.

Rheinwald, J. G. and Green, H. (1975). Serial cultivation of strains of human epidermal keratinocytes: The formation of keratinising colonies from single cells. *Cell 6,* 331-344.

Riddell, R. J., Clothier, R. H. and Balls, M. (1986). Evaluation of three in vitro cytotoxicity assays. *Food and Chemical Toxicology 24*, 469-471.

Rubin, A. L. and Rice, R. H. (1986). Differential regulation by retinoic acid and calcium of transgultaminases in cultured neoplastic and normal human keratinocytes. *Cancer Research 46*, 2356-2361.

Shaw, A. J., Clothier, R. H. and Balls, M. (1990). Loss of transepithelial impermeability of a confluent monolayer of ocular irritancy potential. *ATLA 18*, 145-151.

Shaw, A. J., Balls, M., Clothier, R. H. and Bateman, N. D. (1991). Predicting ocular irritancy and recovery from injury using madin-darby canine kidney cells. *Toxicology in vitro 5/6*, 569-571.

Spielmann, H., Balls, M., Dupuis, J. et al. (1998). The international EU/COLIPA in vitro photoirritancy validation study: Results of Phase II (Blind trial), Part 1: the 3T3 NRU photoirritation test. *Toxicology in Vitro 12*, 305-327.

Strickland, J. A., Stokes, W. S., Casati, S. et al. (2003). Design of a validation study to evaluate in vitro cytotoxicity assays for predicting rodent and human acute systemic toxicity. *Toxicologist 72 (S-1), 157*, Abstract 761.

Suuronen, E. J., Nakamura, M., Watsky, M. A. et al. (2004). Innervated human corneal equivalents as in vitro models for nerve-target cell interactions. *FASEB J. 18*, 170-172.

Ward, R. K., Nation, P. N., Maxwell, M. et al. (1997). Evaluation in vitro of Epidermal cell keratinization. *Toxicology in Vitro 11*, 633-636.

Wilkinson, P. J. and Clothier, R. H. (2005). Comparison of an Animal Product Free with normal growth supplement on growth and barrier integrity of a human corneal epithelial cell line. *ATLA 33*, 1-10.

Woodward, R. M., Wallace, V. P., Arnone, D. D. et al. (2003). Terahertz pulsed imaging of skin cancer in the time frequency domain. *Journal of Biological Physics 29*, 257-261.

Acknowledgments

This work was supported by the sponsors of the FRAME research programme, EU grant. The human corneal cells were kindly donated by Prof. Araki-Sasaki (Japan).

Correspondence to

Dr Richard Clothier
Reader in Cell Toxicology
FRAME Alternatives Laboratory
School of Biomedical Sciences
University of Nottingham, NG7 2UH, UK
e-mail: Richard.Clothier@nottingham.ac.uk

Engineering Human Hepatoma Cells with Key Transcription Factors to Generate Metabolically Competent Hepatic Models

Celia P. Martínez-Jiménez[1], María. J. Gómez-Lechón[1], José V. Castell[1,2] and Ramiro Jover[1,2]

[1]Unidad de Hepatología Experimental, Centro de Investigación, Hospital Universitario La Fe and [2]Departamento de Bioquímica, Facultad de Medicina, Universidad de Valencia, Valencia, Spain.

Summary

Because of the intrinsic variability and limited accessibility of primary cultured human hepatocytes, other cell models (e.g. cell lines) have been considered in drug metabolism studies. Unfortunately, hepatic cell lines express only marginal levels of drug-metabolising cytochrome P450s (CYPs) and are not a real alternative. Lack of CYP expression in hepatomas appears to be the consequence of an altered expression of liver-enriched transcription factors and co-regulators. We have attempted to up-regulate CYP expression by transfecting hepatoma cells with selected factors. Re-expression of one or several of these factors proved to be effective in re-activating different CYPs, and supports the notion that this can be an appropriate strategy to achieve metabolically competent cell lines.

Keywords: cytochrome P450, drug metabolism, hepatoma cells, redifferentiation, transcription factors

Introduction

Lipophilic xenobiotics (e.g. drugs and chemicals), once taken in and absorbed, would tend to accumulate in the body because of their poor water solubility. Biotransformation is an acquired evolution feature by which such xenobiotics are rendered more hydrophilic and, hence, are excreted more easily. This implies that xenobiotics undergo a series of chemical modifications before they can be effectively eliminated from the body.

CYP enzymes are major players in xenobiotics biotransformation, which takes place mainly in the liver (Watkins, 1990; Guengerich, 1993). There are several key CYP genes in the human genome encoding for monooxigenase haemoproteins, which confer xenobiotic metabolism capability, namely, CYP1A2, CYP2A6, CYP2B6, CYP2Cs, CYP2D6, CYP2E1, CYP3A4 and CYP3A5 (Gonzalez, 1990).

Drug metabolism plays a determinant role in drug efficacy, as it governs the concentration of the drug at the site of action. Low or inconsistent metabolism rates result in an inadequate or variable clinical response to the drug, which frequently compromises its therapeutic use. Drug metabolism is the major determinant of drug clearance and the factor that is most frequently responsible for the interindividual differences in drug pharmacokinetics and pharmacological action (Spatzenegger and Jaeger, 1995). Gaining knowledge on the metabolism of a given drug, the enzymes involved, and their inhibition or induction potential is a necessary step in the pharmaceutical development of new pharmaceutical compounds.

Laboratory animals are used extensively to support the research and development of new drugs. However, in addition to ethical issues, it has been proven that experimental animals are not fully reliable predictors of drug metabolism for humans. Therefore, replacement of laboratory animals by alternative, human-derived *in vitro* models is a goal in human pharmacology and toxicology.

Two different *in vitro* approaches are currently being used for the investigation of the metabolism of a drug: a) subcellular fractions (liver microsomes, purified enzymes) and b) metabolically competent cellular models (Castell and Gómez-Lechón, 1997). Liver microsomes are the simplest and most straightforward approach to examining the metabolic stability of a compound and to gaining an overview of its metabolism. Their major limitation is that they contain almost exclusively Phase I activities (i.e. CYPs) and very few Phase II enzymes (i.e. conjugating enzymes). Moreover, incubations cannot be prolonged for more than 1-2 hours, which restricts their use with poorly metabolised compounds (Rodrigues, 1999). The involvement of biotransformation enzymes in the metabolism of a drug can also be easily assessed by incubating the compound with cells that have been genetically manipulated to overexpress a single biotransformation enzyme. These cells have become the tool of choice to ascertain whether a given enzyme is or is not involved in the effective metabolism of a drug and in the production of a given metabolite (Rodrigues, 1999).

Primary human hepatocytes are fully competent metabolic cells, as they retain the expression of both Phase I and II enzymes for several days in culture (Gomez-Lechon and Castell, 2000; Gomez-Lechon et al., 2003; Gomez-Lechon et al., 2004). The cells continuously produce the required cofactors to ensure Phase I and II reactions take place and give a more realistic overview of the metabolic profile of the compound. In addition, these cells respond to enzyme inducers, a phenomenon that can only be followed in a system capable of transcribing and translating CYP genes; that is to say, where enzyme activity and/or

mRNA changes can be properly monitored. Human hepatocyte cultures, because they express the majority of drug metabolising enzymes and are capable of generating a metabolic profile of a drug similar to that found *in vivo*, are considered the gold standard. There is considerable coincidence between the *in vitro* and *in vivo* metabolism of hepatocytes (Ponsoda et al., 2001). However, the restricted accessibility to suitable liver samples has greatly hindered the widespread use of primary cultures of human hepatocytes. Moreover, because of the very limited ability of differentiated hepatocytes to grow *in vitro*, cell cultures need to be prepared each time from liver tissue, which makes it even more difficult to use for routine testing.

Different alternatives have been explored in recent years to overcome such limitations, including immortalisation of adult or foetal human hepatic cells by means of transformation with tumour virus genes, oncogenes or telomerase, conditionally immortalised hepatocytes and cell fusion (Fischbach et al., 1991; Utesch et al., 1992; Pfeifer et al., 1993; Kobayashi et al., 2001; Smalley et al., 2001; Wege et al., 2003). Unfortunately, hepatic cell lines, because of their very low/partial CYP expression, do not constitute a *real* alternative to primary cultured hepatocytes for drug metabolism studies. New strategies are currently being used to up-regulate the expression of drug metabolising enzymes in these cells.

Why do cell lines not express CYP genes? Causes of the poor biotransformation activity of cell lines

Most drug-metabolising CYPs display a tissue-specific and developmental pattern of expression. Hepatic-specific transcription is accomplished by the concerted action of a reduced number of transcription factors. Although the expression of each individual factor is not restricted to hepatocytes, those required to promote liver-specific gene transcription are expressed simultaneously and at high levels in the liver (liver-enriched transcription factors, LETFs). LETFs undergo major changes during the processes of growth and dedifferentiation, which explains the altered pattern of expression of many adult liver-specific genes, among them CYPs, and the fact that they are highly expressed in the adult liver but are nearly absent in the foetal liver (Cereghini, 1996; Schrem et al., 2002; Schrem et al., 2004).

The low biotransformation activity of hepatic cell lines is likely to be the consequence of a very low CYP gene transcription, resulting in trace levels of mRNA and apoprotein synthesis (Rodriguez-Antona et al., 2002). A first possible explanation for this phenomenon is that the specific transcription factors that control the expression of CYP genes are altered (possibly diminished) in hepatomas. Alternatively, CYP transcription could be inhibited as a consequence of increased levels of transcription repressors/co-repressors or an altered functionality of transcription activators. Finally, chromatin condensation and gene silencing as a result of increased DNA methylation / histone deacetylation and/or a decrease in co-activators with acetyltransferase activity could also play a role.

Strategies to overcome the limitations of hepatic cell lines

Overcoming lack of CYP expression by transfecting cells with key transcription factors

There is cumulative evidence suggesting that CYP genes are largely regulated at the transcriptional level. The detection by RT-PCR of minimal but consistent amounts of CYP mRNAs in hepatoma cells (Rodriguez-Antona et al., 2002) suggests that CYP genes have not been totally silenced in these cells by mechanisms such as extensive methylation and chromatin condensation.

Comparative analysis of four major liver transcription factors between HepG2 cells and human hepatocytes, showed that some of them are poorly expressed in hepatoma cells (i.e. C/EBPα mRNA, ca. 15% of human hepatocytes; HNF3α, 25%; HNF1α, 40%), while others (i.e. HNF4 mRNA) are found at similar levels in both cell types (Jover et al., 1998). C/EBPβ, another important LETF for CYP regulation (Martinez-Jimenez et al., 2005), is largely down-regulated in HepG2 and Hep3B when compared to human hepatocytes (fig. 1). Hence, it is conceivable

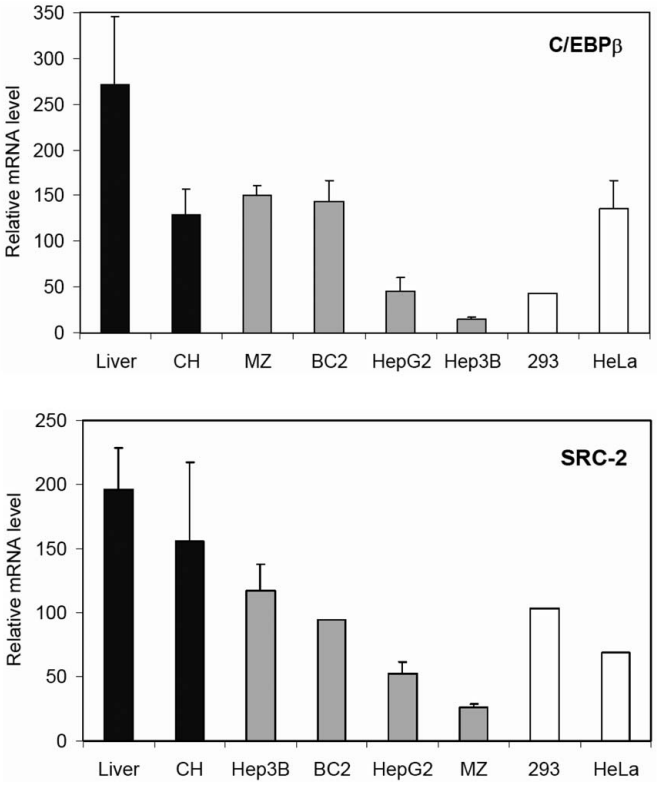


Fig. 1: Expression levels of transcription activators in human hepatoma cells, cultured hepatocytes and liver. Total RNA was purified from human liver, 24 h cultured hepatocytes (CH), hepatoma cells (grey bars) and non-hepatic cell lines (white bars). Specific mRNA levels of the transcription factor C/EBPβ and the coactivator SRC-2 were measured by real-time quantitative RT-PCR. Sample-to-sample variations were normalised by comparison to PBGD cDNA. Bars represent means ± SD (n ≥ 3).

that the lack of CYP gene expression in hepatomas could be the consequence of an altered expression of key regulatory transcription factors.

A first candidate to verify the feasibility of this hypothesis was the CCAAT/enhancer binding protein alpha (C/EBPα), which plays an important role in the differentiation of a number of cell types, among them hepatocytes. C/EBPα expression in the human hepatoma HepG2 was very low in comparison to primary cultured hepatocytes (Jover et al., 1998). The next logical step was to investigate whether restoring levels of this factor in hepatoma cells would have effects on its phenotype. Upon transfection of HepG2 cells with expression vectors encoding for C/EBPα, we found a significant increase in several CYPs of the CYP2 and CYP3 families (Jover et al., 1998; Rodriguez-Antona et al., 2003). This concept was validated with other transcription factors poorly expressed in hepatoma cells: HNF3γ. Three HNF3 isoforms are known in mammals, which are sequentially expressed in embryogenesis. HNF3γ is the only one expressed at the latest stages of endoderm differentiation and is present at high levels in liver tissue (Lai et al., 1993). On the contrary its expression in hepatoma cells is very weak (Lai et al., 1993). Re-expression of HNF3γ caused simultaneous reactivation of several CYPs of the 2C subfamily (Bort et al., 2004). These results definitively suggest that the low expression levels of key transactivators must be relevant in explaining the lack of expression of CYP genes in hepatoma cells.

Although relevant as a *proof of concept*, the above-mentioned results also evidenced the unlikelihood that restoring a single regulatory factor in hepatoma cells could up-regulate all CYP genes. Rather, several factors may be needed for a global reactivation of CYP enzymes in hepatoma cells. The picture that emerges from studies on gene regulation is that transcription factors act cooperatively in a complex cross-regulatory network to determine the hepatic phenotype (Duncan et al., 1998). According to this view it seems reasonable to expect that the re-expression of all major CYP isoforms in hepatoma cells will require the concerted action of several liver-enriched transcription factors.

The experimental strategy to achieve this goal (balanced expression of several key transcription factors) is rather complex. A promising experimental approach is the use of adenoviral expression vectors that show a high efficiency of transfection with no insertational mutagenesis, allow modulation of the level of expression of the transgenes, simultaneous expression of more than one transgene, and are compatible with most mammalian replicating and non-replicating cell lines (Castell et al., 1997). To this end, adenoviruses encoding two relevant liver-enriched transcription factors (HNF3γ and C/EBPα) were successfully generated (Rodriguez-Antona et al., 2003; Bort et al., 2004). HepG2 cells were transduced with various combinations of both adenovirusus, and the impact on CYP expression was assessed. The results of this investigation evidenced that the appropriate combination of both factors, acting synergistically on the endogenous CYP3A4 gene, results in a more than 60-fold increase CYP3A4 mRNA expression (Rodriguez-Antona et al., 2003).

This experimental evidence lends support to the working hypothesis that by an appropriate re-expression of a combination of key transcription factor(s) it is possible to achieve a substantial expression of drug metabolism enzymes in hepatoma cells.

Reactivation of non-functional transcription factors (HNF4α)

HNF4α is a nuclear receptor primarily expressed in liver, gut, kidney and pancreas, which plays a prominent role in the control of liver development, the maintenance of hepatic phenotype and the regulation of many liver-specific genes (glucose, cholesterol, fatty acid metabolism, synthesis of blood coagulation factors) (Sladek et al., 1990). The key role of HNF4α in the hepatic CYP expression was elegantly demonstrated by means of anti-sense technology (Jover et al., 2001).

Surprisingly, HNF4α is present in HepG2 cells at levels comparable to those in hepatocytes, as assessed by RT-PCR and immunoblotting, yet the expression of HNF4-target genes is seriously impaired (Martinez-Jimenez et al., 2004). This raises the question why the HNF4-dependent transactivation is so low in HepG2 cells compared to hepatocytes.

HNF4α exists in several isoforms, all of them capable of binding to the regulatory regions of the genes but with different transactivating potential. An imbalanced expression of HNF4α isoforms could explain why HNF4 is not fully active in HepG2. This does not seem to be the case; the profile of the major HNF4α splicing variants, (α1, α2, α3 and α7) showing a concomitant higher expression in HepG2 than in human hepatocytes (Martinez-Jimenez et al., 2004), cannot account for the loss of function of this transcription factor.

Another possible explanation for the lack of functionality of endogenously expressed HNF4α is that HepG2 cells lack essential specific co-activators. The p160 *steroid receptor co-activator* (SRC) gene family contains three homologous members, which serve as transcriptional co-activators for nuclear receptors and other transcription factors. These co-activators interact with ligand-bound nuclear receptors to recruit histone acetyltransferases to specific enhancer/promoter regions, facilitating chromatin remodelling, assembly of general transcription factors on promoter regions, and transcription of target genes (Xu and Li, 2003). Among the co-activators investigated, SRC-1 and SRC-2 showed a decreased expression level in human hepatoma cells when compared to more differentiated hepatic cells (fig. 1). Interestingly, transfection of SRC-1 in HepG2 cells caused a significant transactivation of the HNF4-regulated CYP2C9 promoter, while it had no effect in HeLa cells, which lack endogenous HNF4α. The experimental data suggest that endogenous HNF4 in HepG2 cells is functional but its activity is limited by the low concentration of SRC co-activators. Re-expression of these limiting co-activators would likely promote the critical interactions and cooperation needed to stimulate HNF4-mediated CYP transcription.

Blocking of transcription repressors

Lack of expression of CYPs in hepatoma cells could be due to an increased level of transcription repressors and co-repressors in these cell lines. For instance, several C/EBP protein isoforms corresponding to full length as well to truncated proteins have been described (Descombes and Schibler, 1991; Ossipow et al.,

1993). The C/EBPβ mRNA, for instance directs production of two isoforms: a 35-kDa LAP (liver-enriched transcriptional activating protein) and a 20-kDa LIP (liver-enriched transcriptional inhibitory protein). Truncated LIP can heterodimerise with the full-length isoform and function as a *dominant negative* regulator, as it lacks most of the transactivation domain but contains the DNA-binding and dimerisation domains (Descombes and Schibler, 1991). According to this, it is feasible that increased expression of LIP isoforms, as it occurs in hepatoma cells, could contribute to the repression of CYP genes.

The chicken ovalbumin upstream promoter-transcription factors (COUP-TFs) are orphan receptors of the nuclear receptor superfamily (Tsai and Tsai, 1997) that were originally characterised as transcriptional activators of the chicken ovalbumin gene but are considered to act as transcription repressors for other nuclear receptors such as HNF4. Direct evidence of the COUP-TF-associated repression of a liver-specific gene regulated by HNF4 is found in the rat ornithine transcarbamylase gene, an ornithine cycle enzyme (Kimura et al., 1993). The levels of some COUP-TF factors are increased in hepatoma cell lines, supporting the hypothesis of this mechanism as a cause for CYP down-regulation in hepatoma cells (fig. 2).

Co-repressors are co-regulators recruited by DNA-bound transcriptional silencers. Co-repressors mediate transcriptional silencing by mechanisms that include direct inhibition of the basal transcription machinery, and recruitment of chromatin-modifying enzymes. The overall picture suggests that transcriptional silencers and co-repressors act in analogy to transcriptional activators and co-activators, but with opposed effects leading to gene silencing (Burke and Baniahmad, 2000).

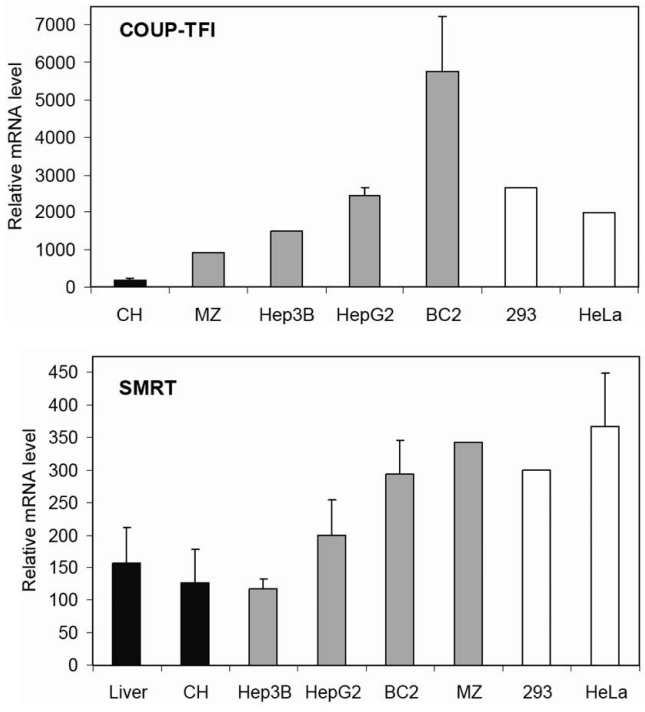


Fig. 2: Expression levels of transcription repressors in human hepatoma cells, cultured hepatocytes and liver. Total RNA was purified from human liver, 24 h cultured hepatocytes (CH), hepatoma cells (grey bars) and non-hepatic cell lines (white bars). Specific mRNA levels of the transcription repressor COUP-TF I and the corepressor SMRT were measured by real-time quantitative RT-PCR. Sample-to-sample variations were normalised by comparison to PBGD cDNA. Bars represent means ± SD (n ≥ 3).

We have examined the levels of two relevant co-repressors, SMRT and NCoR, in hepatoma cells as well as in differentiated adult hepatocytes, and observed significant differences. In some hepatoma cells the levels of SMRT are clearly induced (fig. 2), which is expected to tip the fine-tuned balance between co-activators and co-repressors in hepatoma cells towards a more repressive transcriptional scenario.

Since increased levels of transcription repressors can be an important factor leading to CYP down-regulation in cell lines, blocking expression of these regulators (i.e. siRNA) could be a feasible approach to reactivate CYP expression and metabolism.

Chromatin remodelling: DNA methylation, histone deacetylation and repression of CYP transcription

Cells can modulate gene transcription by DNA methylation and histone deacetylation, both leading to a more compact chromatin structure to which the accessibility of transcriptional activators is impaired.

Methylation of mammalian DNA has long been related to tissue-specific patterns of gene expression (Razin and Riggs, 1980). Methylation of the so-called CpG islands located in and around promoters can lead to chromatin condensation and transcription repression. The mechanism involves methylation of cytosines within 5'-CpG-3' dinucleotides at the pyrimidine ring by DNA methyltransferases, followed by binding of *methyl-binding domain containing proteins* to methylated DNA and subsequent recruitment of co-repressor complexes (e.g. Sin3 and Mi-2/NuRD) (Knoepfler and Eisenman, 1999), which include core histone deacetylases HDAC1 and HDAC2 (Nan et al., 1998). The HDACs remove acetyl groups from the lysine residues found at the N-termini of histones H3 and H4. Their removal results in an increase in the positive charge of the histones, which is thought to condense chromatin by enabling a tighter association between the histones and the negatively charged DNA phosphates. This may in turn silence transcription by assembling methylated sequences into condensed nucleosomes (Nan et al., 1998; Knoepfler and Eisenman, 1999).

For the vast majority of transcription factors, assembly of their response elements into nucleosomes reduces their binding affinity by two to three orders of magnitude. Indeed, the assembly of a TATA box into a nucleosome is a very efficient mechanism to repress transcription. Thus, nucleosomes over the TATA box must be disrupted for the basal transcriptional machinery to function effectively.

An alteration of the DNA methylation/histone acetylation pattern could be responsible for assembling hepatic-specific genes (i.e. CYPs) into a condensed state, rendering the promoter regions inaccessible to the transcription machinery and thus blocking transcription. Indeed, results from our laboratory suggest that chromatin condensation is likely to be an important factor for the low transcription of some CYP genes in hepatoma

cells. For instance, the promoter of CYP2C8 has several putative HNF-3 binding sites; however its mRNA level did not change upon HNF3γ overexpression in HepG2 cells. This can be interpreted as the result of a diminished or hindered accessibility of HNF3γ to its binding sites. Incubation of HepG2 cells overexpressing HNF3γ with Trichostatin A (TSA), a chemical inhibitor of histone deacetylases capable of remodelling chromatin to a transcriptional competent state (Monneret, 2005), resulted in a 20-fold increase of CYP2C8 mRNA levels (Bort et al., 2004). In a similar fashion, transactivation of the endogenous CYP2C19 gene by HNF3γ was considerably improved by TSA, suggesting that chromatin decompactation was also associated with a more effective transactivation of HNF3γ on this gene (Bort et al., 2004).

Conclusions

Hepatoma cell lines show very limited metabolic capacity, due to a very low expression of CYP genes. Low levels of key activating liver-enriched transcription factors (e.g. C/EBPα and HNF3γ) in hepatoma cells have been demonstrated to explain this decreased CYP transcription and function. Tailored re-expression of these missing activators in hepatoma cells causes significant re-activation of relevant CYP genes, thus opening up a new experimental strategy to generate competent cell lines for human drug metabolism studies. Other mechanisms certainly influencing gene regulation could also play a significant role. The existence in hepatoma cells of non-functional CYP regulators, such as HNF4α, specific CYP repressors and chromatin condensation around CYP genes must also be taken into consideration when attempting to develop a drug-metabolism competent cell line.

References

Bort, R., Gomez-Lechon, M. J., Castell, J. V. and Jover, R. (2004). Role of hepatocyte nuclear factor 3gamma in the expression of human CYP2C genes. *Arch. Biochem. Biophys. 426*, 63-72.

Burke, L. J. and Baniahmad, A. (2000). Co-repressors 2000. *Faseb J. 14*, 1876-1888.

Castell, J. and Gómez-Lechón, M. (1997). *In vitro methods in pharmaceutical research*. Academic Press London ISBN 0-12-163390-X.

Castell, J. V., Hernandez, D., Gomez-Foix, A. M. et al. (1997). Adenovirus-mediated gene transfer into human hepatocytes: analysis of the biochemical functionality of transduced cells. *Gene. Ther. 4*, 455-464.

Cereghini, S. (1996). Liver-enriched transcription factors and hepatocyte differentiation. *Faseb J. 10*, 267-282.

Descombes, P. and Schibler, U. (1991). A liver-enriched transcriptional activator protein, LAP, and a transcriptional inhibitory protein, LIP, are translated from the same mRNA. *Cell 67*, 569-579.

Duncan, S., Navas, M., Dufort, D. et al. (1998). Regulation of a transcription factor network required for differentiation and metabolism. *Science 281*, 692-695.

Fischbach, M., Cao, H. W., Diez Ibanez, M. et al. (1991). Maintenance of liver function in long term culture of hepatocytes following in vitro or in vivo Ha-rasEJ transfection. *Cell Biol. Toxicol. 7*, 327-345.

Gomez-Lechon, M. and Castell, J. (2000). Isolation and culture of human hepatocytes. In M. N. Berry and A. M. Edwards (eds). *The hepatocyte review* (11-17). London: Kluwer Academic Publishers.

Gomez-Lechon, M. J., Donato, M. T., Castell, J. V. and Jover, R. (2003). Human hepatocytes as a tool for studying toxicity and drug metabolism. *Curr. Drug. Metab. 4*, 292-312.

Gomez-Lechon, M. J., Donato, M. T., Castell, J. V. and Jover, R. (2004). Human hepatocytes in primary culture: the choice to investigate drug metabolism in man. *Curr. Drug Metab. 5*, 443-462.

Gonzalez, F. J. (1990). Molecular genetics of the P-450 superfamily. *Pharmacol.Ther. 45*, 1-38.

Guengerich, F. (1993). Cytochrome P450 enzymes. *American Scientist 81*, 440-447.

Jover, R., Bort, R., Gomez-Lechon, M. J. and Castell, J. V. (1998). Re-expression of C/EBP alpha induces CYP2B6, CYP2C9 and CYP2D6 genes in HepG2 cells. *FEBS Lett. 431*, 227-230.

Jover, R., Bort, R., Gomez-Lechon, M. J. and Castell, J. V. (2001). Cytochrome P450 regulation by hepatocyte nuclear factor 4 in human hepatocytes: a study using adenovirus-mediated antisense targeting. *Hepatology 33*, 668-675.

Kimura, A., Nishiyori, A., Murakami, T. et al. (1993). Chicken ovalbumin upstream promoter-transcription factor (COUP-TF) represses transcription from the promoter of the gene for ornithine transcarbamylase in a manner antagonistic to hepatocyte nuclear factor-4 (HNF-4). *J. Biol. Chem. 268*, 11125-11133.

Knoepfler, P. S. and Eisenman, R. N. (1999). Sin meets NuRD and other tails of repression. *Cell 99*, 447-450.

Kobayashi, N., Noguchi, H., Westerman, K. A. et al. (2001). Cre/loxP-based reversible immortalization of human hepatocytes. *Cell Transplant. 10*, 383-386.

Lai, E., Clark, K., Burley, S. and Darnell, J. (1993). Hepatocyte nuclear factor 3/fork head or "winged helix" proteins: a family of transcription factors of diverse biologic function. *Proc. Natl. Acad. Sci. USA 90*, 10421-10423.

Martinez-Jimenez, C. P., Gomez-Lechon, M. J., Castell, J. V. and Jover, R. (2005). Transcriptional Regulation of the Human Hepatic CYP3A4: Identification of a New Distal Enhancer Region Responsive to CCAAT/Enhancer-Binding Protein {beta} Isoforms (Liver Activating Protein and Liver Inhibitory Protein). *Mol. Pharmacol. 67*, 2088-2101.

Martinez-Jimenez, C. P., Jover, R., Gomez-Lechon, M. J. and Castell, J. V. (2004). Can Hepatoma Cell Lines be Redifferentiated to be Used in Drug Metabolism Studies? *Altern. Lab. Anim. 32,* 65-74.

Monneret, C. (2005). Histone deacetylase inhibitors. *Eur. J. Med. Chem. 40*, 1-13.

Nan, X., Ng, H., Johnson, C., Laherty, C., et al. (1998). Transcriptional repression by the methyl-CpG-binding protein

MeCP2 involves a histone deacetylase complex. *Nature 393*, 386-389.

Ossipow, V., Descombes, P. and Schibler, U. (1993). CCAAT/enhancer-binding protein mRNA is translated into multiple proteins with different transcription activation potentials. *Proc. Natl. Acad. Sci. U S A 90*, 8219-8223.

Pfeifer, A. M., Cole, K. E., Smoot, D. T. et al. (1993). Simian virus 40 large tumor antigen-immortalized normal human liver epithelial cells express hepatocyte characteristics and metabolize chemical carcinogens. *Proc. Natl. Acad. Sci. U S A 90*, 5123-5127.

Ponsoda, X., Pareja, E., Gomez-Lechon, M. J. et al. (2001). Drug biotransformation by human hepatocytes. In vitro/in vivo metabolism by cells from the same donor. *J. Hepatol. 34*, 19-25.

Razin, A. and Riggs, A. (1980). DNA methylation and gene function. *Science 210*, 604-610.

Rodrigues, A. D. (1999). Integrated cytochrome P450 reaction phenotyping: attempting to bridge the gap between cDNA-expressed cytochromes P450 and native human liver microsomes. *Biochem. Pharmacol. 57*, 465-480.

Rodriguez-Antona, C., Bort, R., Jover, R., Tindberg, N., et al. (2003). Transcriptional regulation of human CYP3A4 basal expression by CCAAT enhancer-binding protein alpha and hepatocyte nuclear factor-3 gamma. *Mol. Pharmacol. 63*, 1180-1189.

Rodriguez-Antona, C., Donato, M. T., Boobis, A. et al. (2002). Cytochrome P450 expression in human hepatocytes and hepatoma cell lines: molecular mechanisms that determine lower expression in cultured cells. *Xenobiotica 32*, 505-520.

Schrem, H., Klempnauer, J. and Borlak, J. (2002). Liver-enriched transcription factors in liver function and development. Part I: the hepatocyte nuclear factor network and liver-specific gene expression. *Pharmacol. Rev. 54*, 129-158.

Schrem, H., Klempnauer, J. and Borlak, J. (2004). Liver-enriched transcription factors in liver function and development. Part II: the C/EBPs and D site-binding protein in cell cycle control, carcinogenesis, circadian gene regulation, liver regeneration, apoptosis, and liver-specific gene regulation. *Pharmacol. Rev. 56*, 291-330.

Sladek, F. M., Zhong, W. M., Lai, E. and Darnell, J. E., Jr. (1990). Liver-enriched transcription factor HNF-4 is a novel member of the steroid hormone receptor superfamily. *Genes. Dev. 4*, 2353-2365.

Smalley, M., Leiper, K., Tootle, R. et al. (2001). Immortalization of human hepatocytes by temperature-sensitive SV40 large-T antigen. *In Vitro Cell Dev. Biol. Anim. 37*, 166-168.

Spatzenegger, M. and Jaeger, W. (1995). Clinical importance of hepatic cytochrome P450 in drug metabolism. *Drug Metab. Rev. 27*, 397-417.

Tsai, S. Y. and Tsai, M. J. et al. (1997). Chick ovalbumin upstream promoter-transcription factors (COUP-TFs): coming of age. *Endocr. Rev. 18*, 229-240.

Utesch, D., Arand, M., Thomas, H. (1992). Xenobiotic-metabolizing enzyme activities in hybrid cell lines established by fusion of primary rat liver parenchymal cells with hepatoma cells. *Xenobiotica 22*, 1451-1457.

Watkins, P. (1990). Role of cytochromes P450 in drug metabolism and hepatotoxicity. *Semin Liver Dis 10*, 235-250.

Wege, H., Le, H. T., Chui, M. S et al.. (2003). Telomerase reconstitution immortalizes human fetal hepatocytes without disrupting their differentiation potential. *Gastroenterology 124*, 432-444.

Xu, J. and Li, Q. (2003). Review of the in vivo functions of the p160 steroid receptor coactivator family. *Mol. Endocrinol. 17*, 1681-1692.

Acknowledgements

This research was supported in part by a grant from the Spanish Ministry of Science and Technology (SAF 2003-09353) and from an EU Integrated Research Project (PREDICTOMICS. LSHB-CT-2004-504761).

Correspondence to

Prof. Ramiro Jover
Unidad de Hepatología Experimental
Centro de Investigación
Hospital Universitario La Fe
Avda de Campanar 21
46009 Valencia, Spain
e-mail: Ramiro.Jover@uv.es

Effects of Trichostatin A on Apoptosis-regulating Proteins During Hepatocyte Isolation

Tamara Vanhaecke, Mathieu Vinken, Tom Henkens, Sarah Snykers, Greetje Elaut, Peggy Papeleu and Vera Rogiers*

Department of Toxicology, Vrije Universiteit Brussel, Brussels, Belgium

Summary

In this study, the expression levels of key apoptosis-regulating proteins were investigated by Western blot in primary rat hepatocytes during the various steps of their isolation in the presence and absence of trichostatin A (TSA). It was found that addition of 1 μM TSA to the perfusate used for hepatocyte isolation reduced procaspase-3 cleavage as well as the expression of Apaf-1, the tumour suppressor p53, and the pro-apoptotic Bcl-2 proteins Bid and Bax. These findings indicate that, already during the isolation of primary hepatocytes from liver, TSA might direct cell signalling towards a p53-dependent reduction of apoptosis.

Keywords: trichostatin A, hepatocyte isolation, apoptosis, culture

Introduction

Histone deacetylase (HDAC) inhibitors, with trichostatin A (TSA) as a prototype for hydroxamate-containing molecules, are nowadays substantially promoted as potential drugs against various tumorous and non-tumorous diseases (Vanhaecke et al., 2004a; Elaut et al., 2005). However, the majority of the data in the literature arises from studies with tumour cells and cancerogenous cell lines and only very limited data on primary cells is available. Our research is concerned with the biological effects of HDAC inhibitors on primary parenchymal liver cells. After all, hepatocytes are the most important site of biotransformation within our body and information on the effect of HDAC inhibitors on hepatocytes is important for further drug development of this group of molecules.

Next to the clinical application of HDAC inhibitors, a more fundamental reason exists for our interest in hepatocytes, namely their use as a 'differentiation-inducing agent' in primary cultures of hepatocytes. Indeed, previous research carried out by our group has demonstrated that, independent of the conditions used, cultured primary hepatocytes dedifferentiate as a function of time, thereby loosing their specialised functions including xenobiotic biotransformation (reviewed in Vanhaecke et al., 2005 and Papeleu et al., 2002). This dedifferentiation process results from a cell cycle re-entry already induced during isolation of the hepatocytes from the liver (Loyer et al., 1996a, 1996b; Michalopoulos and DeFrances, 1997). We could clearly show that the induction of cell cycle arrest by the addition of HDAC inhibitors to the culture medium of primary hepatocytes not only prolonged the lifespan of the cultures by delaying apoptosis (Papeleu et al., 2003; Vanhaecke et al., 2004b), but also enhanced their differentiated state. This was evidenced by increased albumin secretion, increased expression of cytochrome P450 and liver-enriched transcription factors, and improved gap junctional intercellular communication (Rogiers et al., 2004).

A common observation, though, when using HDAC inhibitors in primary hepatocyte cultures, is that more pronounced effects are observed with respect to cell cycle signalling (Papeleu et al., 2003) and gap junctional intercellular communication (Vinken et al., 2005) when exposure to the compound already starts during hepatocyte isolation compared to its addition at the time of cell plating. In the present work, we have extended this finding with regard to apoptosis by investigating the expression of a number of key apoptosis-regulating proteins during the hepatocyte isolation procedure in the presence and absence of TSA.

Materials and methods

Chemicals

Crude collagenase type I, bovine serum albumin fraction V, R-(+)-TSA, ethylene diamine-N,N,N',N'-tetra acetic acid disodium salt hydrate (EDTA), ethylene glycol-bis(2-aminoethylether)-N,N,N',N'-tetra acetic acid (EGTA), N-(2-hydroxyethyl)piperazine-N'-(2-ethanesulfonic acid) (HEPES), dithiothreitol (DDT), poly-oxy-ethylene sorbitan monolaurate (Tween-20), and phenylmethylsulfonylfluoride (PMSF) were purchased from Sigma-Aldrich. All other chemicals were readily available commercial products of analytical grade and were used without further purification.

Rat hepatocyte isolation and sampling

Rat hepatocytes were isolated (Papeleu et al., 2005a), either in the presence or absence of 1 μM TSA, from outbred adult male Sprague-Dawley rats (200-250 g; Charles River Laboratories, Brussels, Belgium), which were kept under controlled environ-

* Tamara Vanhaecke is a postdoctoral research fellow of the Fund for Scientific Research Flanders (FWO-Vlaanderen) Belgium.

mental conditions (12 h light-dark cycle) and fed a standard diet (Animalabo A 04) with water *ad libitum*. Procedures for housing of the animals and for isolation of the rat hepatocytes were approved by the local ethical committee of the Vrije Universiteit Brussel (VUB, Brussels, Belgium). Cell suspension aliquots were taken during the various steps of the hepatocyte isolation procedure, i.e. upon dissociation (1), filtration (2), first wash (3), second wash (4) and obtaining the freshly isolated hepatocytes (5).

Preparation of cell lysates

Cell suspensions were centrifuged at 2376 g for 5 minutes at 4°C and washed twice in ice-cold PBS. They were resuspended in 50-100 µl lysis buffer with pH 7.5 (50 mM HEPES, 15 mM NaCl, 1 mM EDTA, 2.5 mM EGTA, 5 mM DDT, 0.1 mM PMFS, 0.1% (v/v) Tween, 10% (v/v) glycerol and 1x EDTA-free protease inhibitors cocktail (Roche)) and sonicated on ice (15 pulses). After one hour on ice, the lysates were centrifuged at 13791 g for 5 minutes at 4°C and protein concentrations were determined using a Bradford protein assay kit (Bio-Rad) with bovine serum albumin as standard.

Western blot analysis

25 µg of total cellular protein was loaded onto 12% SDS-PAGE and transferred onto nitrocellulose membranes. Equal protein loading was controlled by reversible 0.1% Ponceau Red staining of the membranes. After blocking the membranes for 1 hour in 0.1% Tween-20 containing 5% (w/v) non-fat milk, the blots were probed overnight at 4°C with primary antibody, followed by a one hour incubation at room temperature with appropriate horseradish peroxidase-conjugated secondary antibody. Antigen-antibody complexes were visualised by enhanced chemiluminescence (Super Signal West Pico©). The primary antibodies used in this study were reactive to (pro)caspase-3 (rabbit; Calbiochem), Bax (mouse; Santa Cruz Biotechnology), Bid (rabbit; R&D Systems), Apaf-1 (mouse, Abcam), Survivin (mouse, Abcam), p53 (mouse, BD Pharmingen) and GAPDH (mouse, Abcam). The goat anti-mouse and goat anti-rabbit secondary antibodies were from Dako Diagnostics.

Results and discussion

Previous research carried out by our team has clearly shown that, under various culture conditions, TSA delays the onset of spontaneous apoptosis in hepatocytes as evidenced by decreased activation of caspase-3 and caspase-8, downregulation of Bid and upregulation of BclxL (Papeleu et al., 2003; Vanhaecke et al., 2004b). An even more pronounced inhibition of apoptosis in hepatocyte cultures has been observed with the more stable HDAC inhibitor ω-carboxypentyl p-dimethylaminobenzamide hydroxamate (Papeleu et al., 2005b). The anti-apoptotic properties of both hydroxamate-based HDAC inhibitors within primary hepatocytes stand in major contrast to the pro-apoptotic properties observed in hepatoma cell lines (Herold et al., 2002), suggesting an important difference in response towards "healthy" primary cells and their tumorous counterparts. Recently, this discrepancy was confirmed by another group for the HDAC inhibitors valproate and ITF2357 (Armeanu et al., 2005). This supports the belief that HDAC inhibitors exhibit a specific selectivity towards tumour cells, which explains the existing interest in their further development as cytostatics.

The biological outcome elicited by TSA in primary hepatocyte cultures depends, however, on the onset of exposure to the compound. As such, TSA induces differential cell cycle arrest when added at the time of cell plating (S phase) or during perfusion of the liver (early G1 phase) (Papeleu et al., 2003). Likewise, gap junctional intercellular communication is significantly more improved during culture when TSA treatment is started during hepatocyte isolation and continued thereafter (Vinken et al., 2005). Here, we investigated whether the previously observed TSA-mediated inhibition of apoptosis in hepatocyte cultures (Papeleu et al., 2003; Vanhaecke et al., 2004b) is already initiated during the hepatocyte isolation procedure. From figure 1 it is clear that TSA downregulates procaspase-3 (34kDa) during the initial steps of the isolation procedure. In addition, the lower expression levels of the large subunit (p18 fragment) in cells perfused with TSA point to reduced caspase-3 cleavage and thus less apoptosis in hepatocytes isolated in the

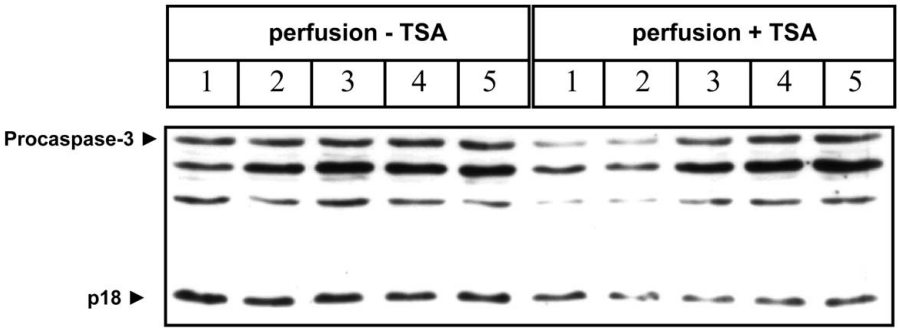


Fig. 1: Effect of TSA on caspase-3 activation during the isolation of primary rat hepatocytes.
Samples were taken during the various steps of hepatocyte isolation, i.e. 1: dissociation, 2: filtration, 3: first wash, 4: second wash, 5: freshly isolated hepatocytes. Cellular extracts were probed for caspase-3 by Western blot analysis. Equal loading was confirmed by Ponceau Red staining.

presence of TSA. This is further evidenced by a decreased expression of Apaf-1 and the pro-apoptotic Bcl-2 proteins Bax and Bid in TSA-exposed cells compared to non-exposed cells (see fig. 2). Moreover, a clear downregulation of the tumour suppressor p53, known to have a direct influence on the expression of Bid, Bax and Apaf-1 (Hofseth et al., 2004), is also observed in hepatocytes isolated in the presence of TSA. This indicates that, within primary hepatocytes, TSA probably inhibits mitochondrial-mediated apoptosis through a decreased p53 expression. Since p53 is also responsible for the direct activation of p21, which in turn leads to a cell cycle arrest in the G1/S phase (Hofseth et al., 2004), the presently observed TSA-mediated downregulation of p53 might also explain the fact that the cell cycle arrests seen in primary hepatocytes are independent of p21 (Papeleu et al., 2003).

Conclusions

The HDAC inhibitor TSA is known to interfere with proliferation, differentiation and apoptosis in a variety of tumour and non-tumour cell lines, including hepatoma cells. However, studies performed in our lab have clearly shown that, in contrast to hepatoma cells, TSA inhibits apoptosis in primary hepatocyte cultures. Here, we demonstrate that this inhibition of apoptosis probably results from a downregulation of p53, initiated during the isolation of the hepatocytes from the liver. As such, these results may contribute to the establishment of a new, mechanistically-based method to acquire and maintain well-differentiated and more viable primary hepatocytes in culture.

References

Armeanu, S., Pathil, A., Venturelli, S. et al. (2005). Apoptosis on hepatoma cells but not on primary hepatocytes by histone deacetylase inhibitors valproate and ITF2357. *J. Hepatol. 42*, 210-217.

Elaut, G., Rogiers, V. and Vanhaecke, T. (2005). The pharmaceutical potential of histone deacetylase inhibitors. *Curr. Pharm. Design*, submitted.

Herold, C., Ganslmayer, M., Ocker, M. et al. (2002). The histone-deacetylase inhibitor Trichostatin A blocks proliferation and triggers apoptotic programs in hepatoma cells. *J. Hepatol. 36*, 233-240.

Hoffman, W. H., Biade, S., Zilfou, J. T. et al. (2002). Transcriptional repression of the anti-apoptotic survivin gene by wild type p53. *J. Biol. Chem. 277*, 3247-3257.

Hofseth, L. J., Perwez Hussain, S. and Harris, C. C. (2004). p53: 25 years after its discovery. *Trends Pharmacol. Sci. 25*, 177-181.

Johnson, A. L., Langer, J. S. and Bridgham, J. T. (2002). Survivin as a cell cycle-related and antiapoptotic protein in granulose cells. *Endocrinol. 143*, 3405-3413.

Loyer, P., Cariou, S., Glaise, D. et al. (1996b). Growth factor dependence of progression through G1 and S phases of adult rat hepatocytes in vitro. *J. Biol. Chem. 271*, 11484-11492.

Loyer, P., Glaise, D., Cariou, S. et al. (1996a). Expression and activation of cdks (1 and 2) and cyclins in the cell cycle progression during liver regeneration. *J. Biol. Chem. 269*, 2491-2500.

Michalopoulos, G. K. and DeFrances, M. C. (1997). Liver regeneration. *Science 276*, 60-66.

Papeleu, P., Loyer, P., Vanhaecke, T. et al. (2003). Trichostatin A induces differential cell cycle arrests but does not induce apoptosis in primary cultures of mitogen-stimulated rat hepatocytes. *J. Hepatol. 39*, 374-382.

Papeleu, P., Coutant, A., Elaut, G. et al. (2005b). The histone deacetylase inhibitor Me_2N-BAVAH induces an early G1 cell cycle arrest and prevents spontaneous apoptosis in primary rat heptocytes. *J. Hepatol.*, submitted.

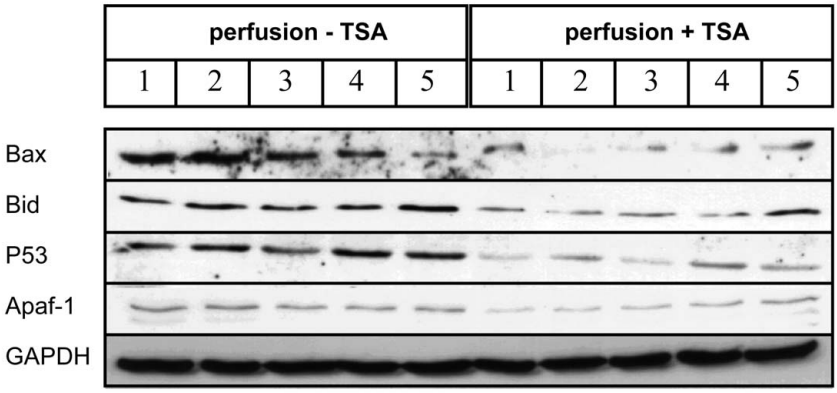


Fig. 2: Effects of TSA on the expression levels of Bid, Bax, p53 and Apaf-1 during the isolation of primary rat hepatocytes.
Samples were taken during the various steps of hepatocyte isolation, i.e. 1: dissociation, 2: filtration, 3: first wash, 4: second wash, 5: freshly isolated hepatocytes. Cellular extracts were probed for the various apoptosis-regulating proteins by Western blot analysis. Equal loading was confirmed with Ponceau Red staining and expression of the housekeeping gene GAPDH.

Papeleu, P., Elaut, G., Rogiers, V. and Vanhaecke, T. (2002). Cell cultures as in vitro tools for biotransformation studies. In S. G. Pandalai (ed.), *Recent Research Developments in Drug Metabolism and Disposition* (199-234). Kerala: Transworld Research Network.

Papeleu, P., Vanhaecke, T., Henkens, T. et al. (2005a). Isolation of rat hepatocytes. In I. R. Philips and E. A. Shepard (eds.), *Cytochrome P450 Protocols. Series in: Methods in Molecular Biology* (229-237). Totowa: Human Press Inc.

Rogiers, V., Snykers, S., Papeleu, P. et al. (2004). Differentiation of stem cells and stabilization of phenotypical properties of primary cells. *PCT/EP2004/012134*, 1/11/2004-1/11/2005, Vrije Universiteit Brussel, Department of Toxicology.

Vanhaecke, T. and Rogiers, V. (2005). Hepatocyte cultures in drug metabolism and toxicological research and testing. In I. R. Philips and E. A. Shepard (eds.), *Cytochrome P450 Protocols. Series in: Methods in Molecular Biology* (209-227). Totowa: Human Press Inc.

Vanhaecke, T., Henkens, T., Kass, G. E. N. and Rogiers, V. (2004b). Effect of the histone deacetylase inhibitor Trichostatin A on spontaneous apoptosis in various types of adult rat hepatocyte cultures. *Biochem. Pharmacol. 68*, 753-760.

Vanhaecke, T., Papeleu, P., Elaut, G. and Rogiers, V. (2004a). Trichostatin A-like hydroxamate histone deacetylase inhibitors as therapeutic agents: toxicological point of view. *Curr. Med. Chem. 11*, 1629-1643.

Vinken, M., Henkens, T., Vanhaecke, T. et al. (2005). Trichostatin A triggers gap junctional intercellular communication in primary cultures of adult rat hepatocytes. *Toxicol. Sci.*, submitted.

Acknowledgements

We gratefully express our thanks to Mrs G. De Pauw, Mrs S. Coppens, Mrs H. Mertens, Mr B. Degreef and Mrs M. Strucelj for their technical assistance. This work was supported by the Fund for Scientific Research - Flanders (FWO), the Research Council (OZR) of the Vrije Universiteit Brussel, Belgium, and the European STREP-project Predictomics no. 504761.

Correspondence to

Dr. Tamara Vanhaecke
Vrije Universiteit Brussel
Department of Toxicology, Dermato-cosmetology and Pharmacognosy
Laarbeeklaan 103
1090 Brussels
Belgium
e-mail: Tamara.Vanhaecke@vub.ac.be

Session 7.6
Non-genotoxic carcinogenicity: Mechanistic perspectives for alternatives

Cancerous Contradictions: The Mis-Regulation of Human Carcinogens Based on Animal Data*

Andrew Knight[1], Jarrod Bailey[2] and Jonathan Balcombe[3]
[1]Animal Consultants International, London, UK; [2]Faculty of Medical Sciences, University of Newcastle upon Tyne, UK;
[3]Physicians Committee for Responsible Medicine, Washington DC, USA

Summary
The regulation of human exposures to potential carcinogens constitutes society's most important use of animal carcinogenicity data. However, for environmental contaminants of greatest U.S. concern, we found that in most cases (58.1%; 93/160) the U.S. Environmental Protection Agency (EPA) considered the animal data inadequate to support a classification of probable human carcinogen or non-carcinogen.
The World Health Organisation's International Agency for Research on Cancer (IARC) is a leading international authority on carcinogenicity assessments. For chemicals lacking human exposure data (the great majority), IARC classifications of identical chemicals were significantly more conservative than EPA classifications ($p<0.0001$), indicating that: (i) the EPA is over-reliant on animal carcinogenicity data, (ii) as a result, it tends to over-predict carcinogenic risk, and (iii) the true predictivity for human carcinogenicity of animal data is even poorer than indicated by EPA figures alone. EPA policy erroneously assuming that tumours in animals are indicative of human carcinogenicity is implicated as the greatest source of these errors.

Keywords: animal experiment, animal test, bioassay, cancer prevention, carcinogenicity, chemical classification, chemical safety, risk assessment

Introduction

Since the first animal carcinogenicity test in 1915, when Yamagiwa and Ichikawa showed that coal tar applied to rabbits' ears caused skin carcinomas, several thousand carcinogenicity bioassays have been conducted, with the objective of determining human carcinogenic risks for the great majority of chemicals lacking human exposure data (Huff, 1999). However, animal carcinogenicity testing remains a controversial area of research.

Proponents claim that all known human carcinogens that have been studied in sufficient animal species have produced positive results in one or more species (Wilbourn et al., 1986; Tomatis et al., 1989; Rall, 2000). Critics respond that if enough animal testing is conducted, carcinogenesis will eventually occur in *some* species, regardless of human cancer risk. A study published in *Mutagenesis* found that of 20 human *non*-carcinogens, 19 produced carcinogenic effects in animals (Ennever et al., 1987).

The most important use of animal carcinogenicity data lies in the regulation of human exposures to potential carcinogens. The U.S. Federal agency most responsible for regulating exposures to environmental contaminants is the Environmental Protection Agency (EPA, undated a), and the chemicals of greatest public

* Summarised from the poster *Animal carcinogenicity studies: poor human predictivity* by Andrew Knight, Jarrod Bailey and Jonathan Balcombe, which received the Animal Welfare Poster Award from Deutscher Tierschutzbund (the German Animal Welfare Federation) at the 5th World Congress on Alternatives and Animal Use in the Life Sciences, Berlin, 25 August 2005. Reproduced with permission from the complete paper: Knight, A., Bailey, J., Balcombe, J. (2006). Animal carcinogenicity studies: 1. poor human predictivity. *Alternatives to Laboratory Animals 34(1)*, 19-27.

health concern (EPA, undated b) are listed within its Integrated Risk Information System (IRIS) chemicals database, along with their animal toxicity data and consequent human carcinogenicity assessments (EPA, undated c).

To assess the utility of animal carcinogenicity data in deriving human carcinogenicity assessments, we surveyed the IRIS chemicals database. To assess the reliability of the EPA carcinogenicity assessments obtained from animal test data, we compared them with those of a leading world authority, the World Health Organization's International Agency for Research on Cancer (IARC).

Methods

The 543 chemicals catalogued in the EPA's IRIS chemicals database (as of January 1, 2004; EPA, undated d) were examined to determine the proportion for which the EPA was able to derive classifications of "probable human carcinogen" or "probable human non-carcinogen" based primarily on animal carcinogenicity data. The relatively few classifications of "definite human carcinogen" relied primarily on available human exposure data. The remaining classifications of "unclassifiable" or "possible human carcinogen" were not considered substantially useful for risk assessment or regulatory purposes. They are excluded from the U.S. National Toxicology Program annual *Report on Carcinogens* (NTP, 2002).

Of the 177 chemicals considered by the EPA to possess at least limited human or animal data, 128 were assigned human carcinogenicity classifications by both the EPA and the IARC. Of these 128, 17 were considered by the EPA to possess at least limited human data, while 111 were primarily reliant on animal data.

The consistency of classifications between the EPA and the IARC was examined for these two groups by comparing the carcinogenicity classification proportions within each group by chi-square tests,[1] and also by comparing the individual classifications of the 111 chemicals primarily reliant on animal carcinogenicity data.

Chi-squared tests provide statistical calculations of the probability that two data sets, such as EPA and IARC human carcinogenicity classifications, are samples from the same underlying data population, and that any observed differences are simply due to random sampling variation. Large chi-squared (X^2) values reflect increased probabilities that observed differences are due to real differences in underlying data populations.

Results

EPA human carcinogenicity classifications

Of the 543 chemicals catalogued in the EPA's IRIS chemicals database, 235 had been assigned human carcinogenicity classifications. Of these, 17 were classified as definite (A) or probable (B1) human carcinogens on the basis of their human carcinogenicity data. Of the remaining 218 chemicals lacking even limited human data, 160 were deemed to possess animal carcinogenicity data, primarily sourced from the biomedical literature (B2, C, subset of D, and E; tab. 1).

The human utility of animal carcinogenicity data based on EPA figures

Of the 160 EPA chemicals lacking even limited human data (A or B1) but having animal data (B2, C, subset of D, and E), 64 were considered probable human carcinogens (B2), and three were considered probably not carcinogenic to humans (E). The remaining 93 chemicals were considered possible human carcinogens (C; 40) or unclassifiable as to their human carcinogenicity (D; 53) based on animal data considered inadequate to support a stronger classification (tab. 1).

In sum, of those 160 chemicals lacking even limited human

[1] Chi-squared and two-tailed p values were derived from the online statistical calculators available at www.graphpad.com/quickcalcs/index.cfm.
[2] Confidence interval derived via the modified Wald method described by Agresti et al. (1998) as being more accurate than the so-called "exact" method commonly used.
[3] Chi-squared analysis does not allow comparison when one category lacks any data, hence acrylonitrile, assessed as the only possible human carcinogen by IARC, but as a probable human carcinogen (B1) by the EPA, was excluded, yielding a more conservative result.
[4] To allow chi-squared analysis, methacrylate, assessed as unclassifiable by IARC, but as the only probable human non-carcinogen by the EPA, was excluded, yielding a more conservative result.

Tab. 1: EPA human carcinogenicity classifications of IRIS chemicals

EPA human carcinogenicity classification (with basis for classification)	No. of chemicals	% of total
A: Human carcinogen (convincing human data)	11	4.7
B1: Probable human carcinogen (limited human data)	6	2.6
B2: Probable human carcinogen (sufficient animal data)	64	27.2
C: Possible human carcinogen (animal data inadequate for stronger classification)	40	17.0
D: Unclassifiable (animal data inadequate for stronger classification)	53	22.6
D: Unclassifiable (no animal or human data)	58	24.7
E: Probable human non-carcinogen (sufficient animal data)	3	1.3
TOTAL	**235**	

160 chemicals lacking in human data had received a human carcinogenicity assessment primarily on the basis of their animal data.
Data source: EPA Integrated Risk Information System database, 1 January 2004.

data but having animal data, the EPA considered the animal data inadequate to support the substantially useful classifications of probable human carcinogen or probable human non-carcinogen in the majority of cases (93/160; 58.1%, 95% CI: 50.4-65.5)[2].

Comparison of EPA and IARC human carcinogenicity classifications

Of those 177 chemicals considered by the EPA to possess human or animal data (A, B1, B2, C, D with animal data, or E), 128 were also assessed by the IARC. Of these, 17 were considered by the EPA to possess at least limited human data (A or B1), and the remaining 111 EPA carcinogenicity classifications were primarily reliant on animal data.

For those 17 chemicals considered by the EPA to possess at least limited human data, overall EPA classifications were not found to differ significantly from those predicted by IARC classifications ($p = 0.5896$, $X^2 = 0.291$, 1 df, tab. 2)[3].

However, for those 111 chemicals considered by the EPA to lack even limited human data, but to possess animal data, EPA and IARC classifications were very significantly different overall ($p < 0.0001$, $X^2 = 215.548$, 2 df; fig. 1)[4].

The EPA was much likelier than the IARC to assign carcinogenicity classifications indicative of greater human hazard. The EPA classified 60 chemicals as probable human carcinogens and 51 in all other categories, which was very significantly different from the IARC figures of 12 and 99, respectively ($p < 0.0001$, $X^2 = 215.273$, 1 df). Similar disparities were found for possible human carcinogens ($X^2 = 19.771$, 1 df, $p < 0.0001$) and unclassifiable chemicals ($p < 0.0001$, $X^2 = 24.378$, 1 df).

Tab. 2: IARC classifications of EPA chemicals possessing significant human data (EPA categories A or B1)

Human carcinogenicity classification	EPA	IARC
Human Carcinogen (A)	11	12
Probable Human Carcinogen (B1)	6	4
Possible Human Carcinogen	0	1
Total	**17**	**17**

Data sources: The EPA Integrated Risk Information System database, 1 January 2004, and the IARC Monographs Programme on the Evaluation of Carcinogenic Risks to Humans, Volumes 1-82, 1 January 2004.

Comparison of the individual classifications of these 111 chemicals revealed that 67 (60.4%) were assigned an EPA carcinogenicity classification indicative of greater human hazard, 38 (34.2%) were assigned an equivalent classification, and 6 (5.4%) were assigned a classification indicative of lesser human hazard than the corresponding IARC classification of the same chemical.

Discussion

Based on EPA figures alone, the predictivity of animal carcinogenicity data for human hazard, and hence its utility in deriving substantially useful human carcinogenicity classifications, is

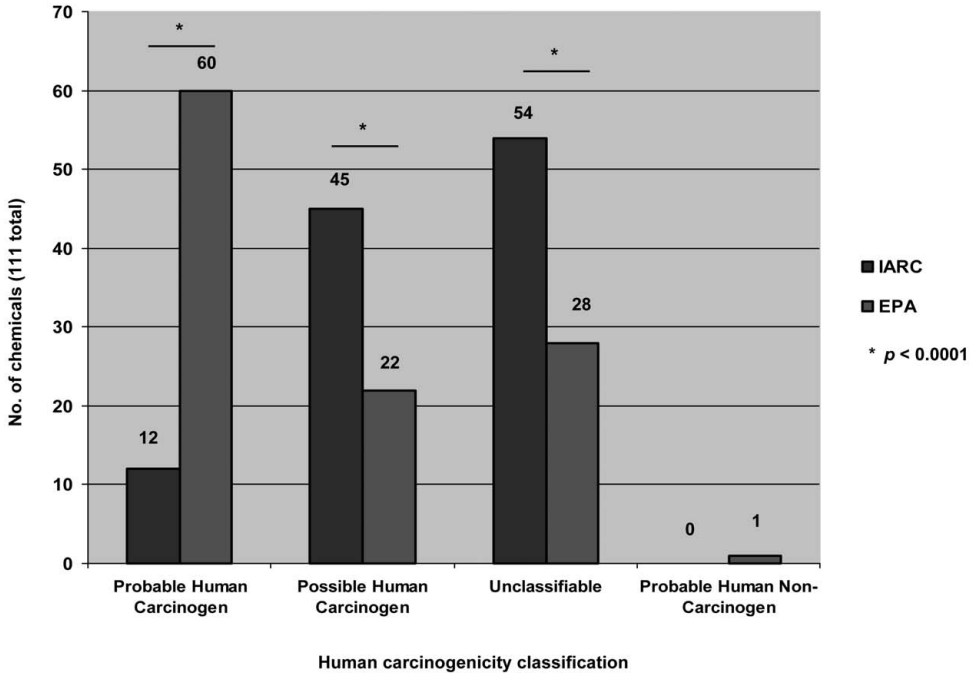


Fig. 1: EPA and IARC human carcinogenicity classifications of chemicals considered by the EPA to lack human data but to possess animal data.
Data source: IARC Monographs Programme on the Evaluation of Carcinogenic Risks to Humans, and the EPA Integrated Risk Information System database, Jan. 1, 2004.

clearly poor. Of those 160 IRIS chemicals lacking even limited human data but possessing animal data, the EPA considered the animal data inadequate to support substantially useful human carcinogenicity classifications in the majority (58%) of cases.

However, IARC assessments of the same chemicals reveal that the human utility of animal carcinogenicity data is probably even lower than indicated by EPA figures. EPA and IARC carcinogenicity classifications were similar only for those chemicals with human data. For those with only animal data, the EPA was much likelier than the IARC to assign carcinogenicity classifications indicative of greater human hazard.

Given that the IARC is recognised as a leading international authority on human carcinogenicity classifications (Tomatis et al., 1993; IARC undated), the very significant differences in classifications of identical chemicals between the IARC and the EPA indicate that:

- (i) in the absence of significant human data the EPA is over-reliant on animal carcinogenicity data,
- (ii) as a result, the EPA tends to over-predict carcinogenic risk, and
- (iii) the true human predictivity for human carcinogenicity of animal data is even poorer than indicated by EPA figures alone.

EPA human carcinogenicity classifications appear to be less scientifically-based than those of the IARC, due to: 1) the varying depth of EPA assessments, due to resource constraints; 2) the less rigorous standards required of data incorporated into EPA assessments; and, in particular, 3) EPA public health-protective policy, which errs on the side of caution by assuming that tumours in animals are indicative of human carcinogenicity (Knight et al. 2006).

Our findings corroborate those of previous investigators. In response to a 2000 Congressional directive, the EPA undertook an evaluation of the data variability and uncertainty within its IRIS assessments. A representative sample of 16 IRIS assessments was subjected to in-depth evaluation by a panel of six independent experts, who concluded that despite being advertised as quantitative science-based classifications, some were, in fact, more grounded in EPA policy favouring classifications indicative of greater human risk (Hogan, 2000).

EPA carcinogenicity assessments may be no more suspect than those of other U.S. regulatory agencies, however. In their survey of 350 representative chemicals, Viscusi and Hakes (1998) found that the carcinogenicity assessments of other U.S. regulatory authorities, particularly the Food and Drug Administration and the Occupational Safety and Health Administration, are even less reflective of actual human risk than those of the EPA. Poor human predictivity of animal carcinogenicity studies was also demonstrated by Tomatis and Wilbourn (1993) and Haseman (2000), and further described by Rall (2000), Ashby and Purchase (1993), Fung et al. (1995) and Ennever and Lave (2003).

Conclusions

By 1998, only about 2,000 (2.7%) of the 75,000 industrial chemicals in use and listed in the EPA's Toxic Substances Control Act inventory had been tested for carcinogenicity (Epstein, 1998). The cost of testing these 2.7% of industrial chemicals was millions of animal lives (Monro et al., 1998; Gold et al., 1999), millions of skilled personnel hours (Gold et al., 1999), and hundreds of millions of dollars (Greek et al., 2000; Stephens et al., 1998).

The most important use of the animal data thus derived is in the regulation of human exposures to potential carcinogens by governmental agencies such as the EPA. However, our results demonstrate that the human predictivity of animal carcinogenicity data was inadequate for the EPA to derive substantially useful human carcinogenicity classifications for the majority (58.1%) of chemicals of greatest public health concern.

Profound differences in human carcinogenicity classifications of identical chemicals between the EPA and the IARC reveal an over-reliance on animal carcinogenicity data by the EPA. The result is that the EPA over-predicts carcinogenic risk. Hence the true human predictivity of animal carcinogenicity data is even poorer than indicated by EPA figures alone.

The sensitivity of the traditional rodent bioassay in detecting human carcinogens for *some* sex-species combinations is not in question. However, its very poor human specificity severely limits its utility for identifying human carcinogens, and its subsequent use in regulating exposures to them. The implementation by regulatory authorities of alternative assays with superior human predictivity is clearly necessary.

References

Agresti, Coull (1998). *The American Statistician 52*, 119-126. The modified Wald method is also described at http://www.graphpad.com/quickcalcs/ConfInterval2.cfm.

Ashby, J. and Purchase, I. F. (1993). Will all chemicals be carcinogenic to rodents when adequately evaluated? *Carcinogenesis 8*, 489-495.

Ennever, F. K. and Lave, L. B. (2003). Implications of the lack of accuracy of the lifetime rodent bioassay for predicting human carcinogenicity. *Regulatory Toxicology and Pharmacology 38*, 52–57.

Ennever, F. K., Noonan, T. J. and Rosenkranz, H. S. (1987). The predictivity of animal bioassays and short-term genotoxicity tests for carcinogenicity and non-carcinogenicity to humans. *Mutagenesis 2(2),* 73-78.

Epstein, S. S. (1998). *The politics of cancer, revisited*. Fremont Center, NY: East Ridge Press.

Fung, V., Barrett, J. and Huff, J. (1995). The carcinogenesis bioassay in perspective: application in identifying human hazards. *Environ. Health Perspect. 103(7-8)*, 680-683.

Gold, L. S., Manley, N. B., Slone, T.H. and Rohrbach, L. (1999). Supplement to the Carcinogenic Potency Database (CPDB): results of animal bioassays published in the general literature in 1993 to 1994 and by the National Toxicology Program in 1995 to 1996. *Environ. Health Perspect. 107 Suppl 4*, 527-600. Based on the total number of animal carcinogenicity tests described.

Greek, C. R. and Greek, J. S. (2000). *Sacred cows and golden geese: the human costs of experiments on animals*. NY: Continuum International.

Haseman, K. (2000). Using the NTP database to assess the value of rodent carcinogenicity studies for determining human cancer risk. *Drug Metab. Rev. 32(2)*, 169-186.

Hogan, K. A. (2000). Characterization of data variability and uncertainty: health effects assessments in the Integrated Risk Information System (IRIS). EPA/635/R-00/005A. National Center for Environmental Assessment, Washington, DC: U.S. Environmental Protection Agency, Office of Research and Development.

Huff, J. (1999). Long-term chemical carcinogenesis bioassays predict human cancer hazards. Issues, controversies, and uncertainties. *Ann. N Y Acad. Sci. 895*, 56-79.

International Agency for Research on Cancer (IARC, undated). IARC monographs programme on the evaluation of carcinogenic risks to humans. http://monographs.iarc.fr, 2004 Jan 1.

Knight, A. and Bailey, J. (2005). Animal carcinogenicity studies: poor human predictivity. *ATLA*, in press.

Knight, A., Bailey, J., Balcombe, J. (2006). Animal carcinogenicity studies: 1. poor human predictivity. *Alternatives to Laboratory Animals 34(1)*, 19-27.

Monro, A. M. and MacDonald, J. S. (1998). Evaluation of the carcinogenic potential of pharmaceuticals. Opportunities arising from the International Conference on Harmonisation. *Drug Saf. 18(5)*, 309-319.

National Toxicology Program (2002). National Toxicology Program Report on Carcinogens, 10th Edn. http://ntp.niehs.nih.gov/ntpweb/index.cfm?objectid=72016262-BDB7-CEBA-FA60E922B18C2540 2005 Jan 13.

Rall, D. P. (2000). Laboratory animal tests and human cancer. *Drug Metab. Rev. 2*, 119-128.

Stephens, M. L., Mendoza, P., Hamilton, T. and Weaver, A. (1998). Unrelieved pain and distress in animals: an analysis of USDA data on experimental procedures. *J. App. Anim. Wel. Sci. 1*, 15-26.

Tomatis, L., Aitio, A., Wilbourn, J. and Shuker, L. (1989). Human carcinogens so far identified. *Jpn. J. Cancer. Res. 80*, 795-807.

Tomatis, L. and Wilbourn, J. (1993). Evaluation of carcinogenic risk to humans: the experience of IARC. In O. H. Iversen (ed.), *New frontiers in cancer causation* (371-387). Washington, D.C: Taylor and Francis.

U.S. Environmental Protection Agency (undated a). http://www.epa.gov, 2004 Jan 29.

U.S. Environmental Protection Agency (undated b). U.S. EPA's process for IRIS assessment development and review. http://www.epa.gov/iris/process.htm, 2003 Dec 10.

U.S. Environmental Protection Agency (undated c). What is IRIS? http://www.epa.gov/iris/intro.htm, 2004 Jan 29.

U.S. Environmental Protection Agency (undated d). IRIS database for risk assessment. http://www.epa.gov/iris/index.html 2004 Jan 1.

Viscusi, W. K. and Hakes, J. K. (1998). Synthetic risks, risk potency, and carcinogen regulation. *J. Policy Anal. Manage. 17(1)*, 52-73.

Wilbourn, J., Haroun, L., Heseltine, E. et al. (1986). Response of experimental animals to human carcinogens: An analysis based upon the IARC Monographs Programme. *Carcinogenesis 7*, 1853-1863.

Acknowledgements

We gratefully acknowledge the assistance of the Physicians Committee for Responsible Medicine (PCRM), Washington DC, in partly funding this research; of the PCRM and the Anti-Vivisection Union of South Australia in funding our attendance at the 5th World Congress on Alternatives and Animal Use in the Life Sciences (5WC); of the Japan Anti-Vivisection Association, Tokyo, in funding our poster based on this paper; and of Deutscher Tierschutzbund (the German Animal Welfare Federation) for their presentation of the Animal Welfare Poster Award for our poster at the 5WC.

Correspondence to

Andrew Knight BSc., BVMS, Cert AW, MRCVS
Director
Animal Consultants International
91 Vanbrugh Court, Wincott St.
London SE11 4NR, UK
e-mail: info@animalconsultants.org
Internet: www.AnimalConsultants.org

Detection of Non-Genotoxic Carcinogens Using Ras-Transfected Bhas 42 Cells

Makoto Umeda
Hatano Research Institute, Food and Drug Safety Center, Kanagawa, Japan

Summary
The cell transformation test employing Bhas 42 cells (v-Ha-ras-transfected BALB/c 3T3 cells) was shown to be a sensitive screening method for the detection of chemicals with different mechanisms of transforming potential. Because of the simplicity of the procedure, it is anticipated that further investigations on Bhas assays will open ways of not only detecting non-genotoxic carcinogens but also of elucidating mechanisms of carcinogenesis.

Keywords: cell transformation assay, Bhas 42 cells, detection of initiating and promoting activities

Introduction

In vitro cell transformation tests can simulate the process of two-stage animal carcinogenesis, and were expected to be an efficient alternative method of animal tumourigenicity tests (IARC/NCI/EPA Working Group, 1985). As the formation of transformed foci is the consequence of the complex process of cell malignisation, the tests can be anticipated to be useful for the detection of not only tumour initiators and promoters, but also of non-genotoxic carcinogens. In spite of this expectation, none of the *in vitro* cell transformation tests have been accepted as routine screening methods, because the tests are thought to be laborious and time-consuming compared with current routine genotoxicity tests.

Ohmori et al. have developed an *in vitro* cell transformation assay for tumour promoters using Bhas 42 cells (Ohmori et al., 2004). The cells are v-Ha-ras-transfected BALB/c 3T3 cells and were considered initiated (Sasaki et al., 1988 and 1990). Recently, in addition to the promotion assay, we found that Bhas cells could be transformed after treatment with initiators, and thus developed an assay method for the evaluation of initiating activity using these cells (Asada et al., 2005). The Bhas transformation assay is a sensitive method and has many advantages, such as shortened experimental time, use of less material and simplicity of the procedure.

Here, we introduce the method for detecting initiating and promoting activities using Bhas 42 cells, and report results obtained by the method, some of which will appear elsewhere (Asada et al., in press; Ohmori et al., 2005).

Materials and methods

Bhas 42 cells were routinely cultured in Minimum Essential Medium (MEM) supplemented with 10% foetal bovine serum (FBS) (M10F). For transformation assays, Dulbecco's modified Eagle's medium/Ham's F12 (DMEM/F12) supplemented with 5% FBS (DF5F) was used. DF2I2F medium used for the growth assay consisted of DMEM/F12 supplemented with 2 µg/ml insulin and 2% FBS.

Experimental procedures are depicted in fig. 1, in which the original BALB/c 3T3 cell transformation assay was included for comparison. In the Bhas 1st-stage assay for the detection of initiating activity, $4x10^3$ cells in 2 ml were seeded into 6-well plates. Test chemicals were added on day 1, and from day 4 the cells were cultured in fresh DF5F medium until day 21. In the Bhas 2nd-stage assay for the detection of promoting activity, $1.4x10^4$ cells were inoculated into wells, and test chemicals were added to the medium on days 4, 7 and 11 when the medium was changed. Cultures were continued until day 21. After fixation and staining, transformed foci were scored under a stereomicroscope. Transformed foci were judged by morphological characteristics: deep basophilicity, dense multilayering of cells, random orientation of cells at the edge of foci, and more than 50 cells within a focus.

Cell growth was assayed using the crystal violet staining method (Saotome et al., 1989). Cells cultured in 24-well plates were fixed with formalin and stained with crystal violet (CV) solution. CV was extracted from the stained cells in each well, and optical density of extracted CV was measured at 540 nm. The results were expressed as percentage of absorbance compared to the solvent control culture.

Results and discussion

Detection of both initiating and promoting activities and tests on PAHs

Typical tumour initiators were examined in Bhas 1st-stage and 2nd-stage assays. Fig. 2 shows results of treatment with typical tumour initiators, *N*-methyl-*N*'-nitro-*N*-nitrosoguanidine (MNNG) and 3-methylcholanthrene (MCA). Induction of transformed foci was observed in the 1st-stage assay, but not in the 2nd-stage assay. In contrast, typical tumour promoters, 12-*O*-

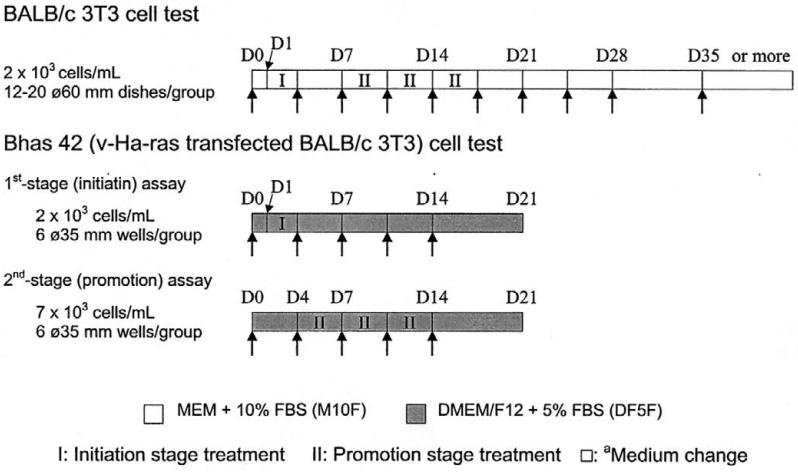


Fig. 1: *In vitro* cell transformation protocols using BALB/c 3T3 cells and Bhas 42 cells

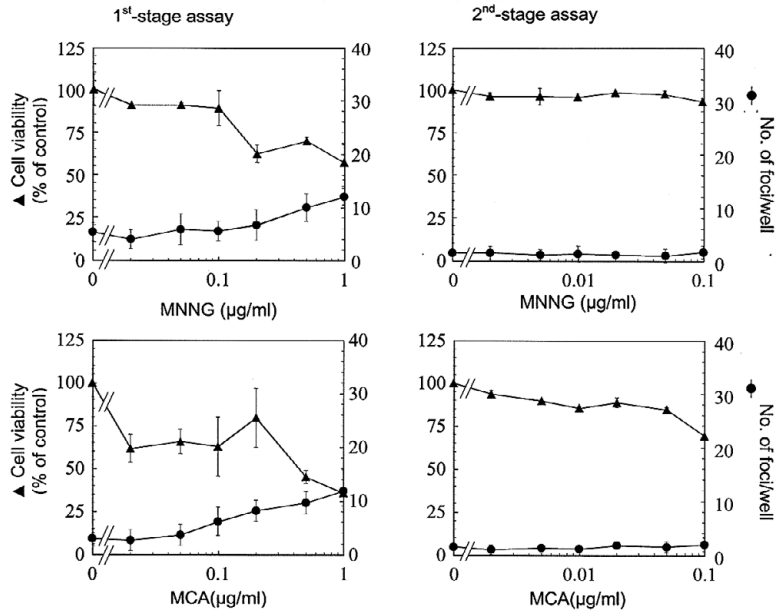


Fig. 2: Bhas 1st-stage and 2nd-stage transformation assays on typical tumour initiators

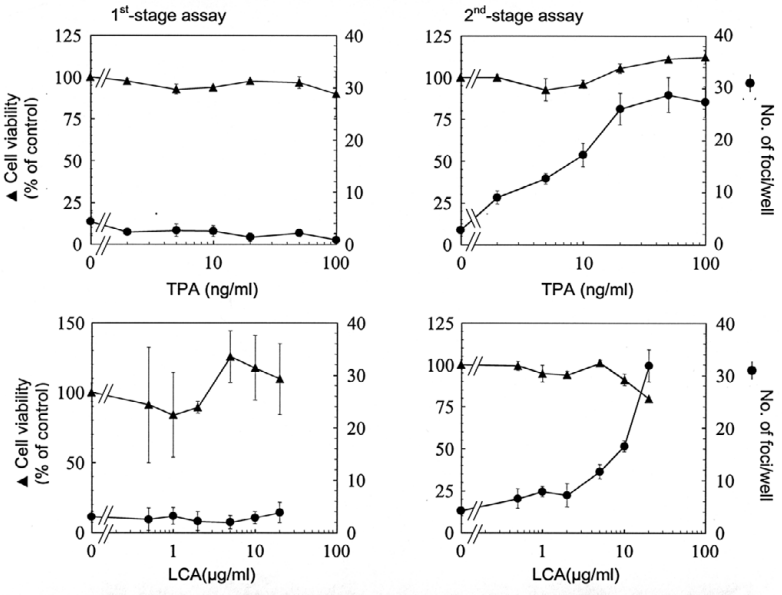


Fig. 3: Bhas 1st-stage and 2nd-stage transformation assays on typical tumour promoters

tetradecanoylphorbol-13-acetate (TPA) and lithocholic acid (LCA), were negative in the 1st-stage assay, but positive in the 2nd-stage assay (see fig. 3).

Fig. 4 shows results of repeated treatment with MCA (100 ng/ml) and/or TPA (50 ng/ml). In this experiment, Bhas 42 cells seeded at $4x10^3$ cells/well were used. Repeated treatment with MCA in the growth phase (1^{st}-stage) increased the number of transformed foci (schedule H), but treatment in the stationary phase (2^{nd}-stage) (schedule D) did not. The opposite situation was observed upon treatment with TPA (schedule E and I). Some tumour initiators are considered to have tumour-promoting activity owing to the experimental observation that repeated application at subtumourigenic doses can induce tumours in mouse skin without any post-treatment with promoter (Berenblum, 1941). Our observation that MCA was effective only when cells were in a dividing phase and TPA was active when cells were in a stationary phase, suggests that the action of tumour initiators is completely different from that of tumour promoters.

Results from both assays on various polycyclic aromatic hydrocarbons (PAHs) and some others are summarised in table 1. Chemicals showing only initiating activity, promoting activity, and both positive and both negative activities are listed. Thus, Bhas cells were reactive to PAHs. The results showing that there are various kinds of PAHs with initiating and/or promoting activities suggest the hazardous nature of mixtures of PAHs such as particulate matter in polluted air, etc.

Collaborative work on Bhas 2^{nd}-stage assay

In collaboration with 14 laboratories in Japan, an inter-laboratory study was conducted in order to validate transferability and applicability of the Bhas 2^{nd}-stage assay (Ohmori et al., 2005). Detailed results of this study were presented in the poster session of this meeting. Briefly, in addition to TPA and LCA, 12

Tab. 1: Summary of Bhas transformation assays on polycyclic aromatic hydrocarbons and typical initiators and promoters

Result of 1st stage assay	Result of 2nd-stage assay Positive	Negative
Positive	Benz[a]anthracene	3-Methylcholanthrene
	Benz[b]anthracene	*N*-Methyl-*N*'-nitro-*N*-nitrosoguanidine
	Chrysene	Aflatoxin B_1
	Perylene	Benzo[a]pyrene
		7,12-Dimethylbenz[a]anthracene
Negative	12-*O*-Tetradecanoylphorbol-13-acetate	anthracene
	Lithocholic acid	Acenaphthylene
	Okadaic acid	Coronene
	Benzo[e]pyrene	9,10-Diphenylanthracene
	Benzo[ghi]perylene	Naphthalene
	1-Nitropyrene	Phenanthrene
	Pyrene	

Tab. 2: Results of collaborative studies of Bhas 2^{nd}-stage assay on 14 chemicals

Chemical	No. of Lab. examined	Judgment			Type of promoter
		Positive	Equivocal	Negative	
TPA	12	12	0	0	T-type
LCA	12	12	0	0	L-type
Anthraline	4	0	2	2	-
Catechol	4	3	0	1	M-type
Dexamethasone	4	1	1	2	-
Diethylstilbestrol	4	0	0	4	-
17β-Estradiol	4	0	0	4	-
Insulin	4	4	0	0	T-type
Mezerein	4	4	0	0	T-type
Okadaic acid	4	3	0	1	L-type
PDD	4	4	0	0	T-type
4α-Phorbol	4	0	0	4	-
Progesterone	4	4	0	0	M-type
Sodium saccharin	4	4	0	0	M-type

chemicals were assayed under blinded conditions. Each chemical was examined in 4 laboratories. Results of 3 positive chemicals are representatively shown in fig. 5. Judgments obtained from the different laboratories were highly consistent (see tab. 2).

This study suggested that there are at least three types of chemicals inducing positive results in the 2nd-stage assay (see fig. 6). Those designated as T(TPA)-type induced extreme growth enhancement, and included TPA, mezerein, PDD and insulin. LCA and okadaic acid belonged to the L(LCA)-type category, in which transformed foci were induced at concentrations showing growth inhibition. In contrast, progesterone, catechol and sodium saccharin (M(miscellaneous)-type) induced foci at concentrations with little to slight growth inhibition. Categorisation of these types is important to decide on doses to

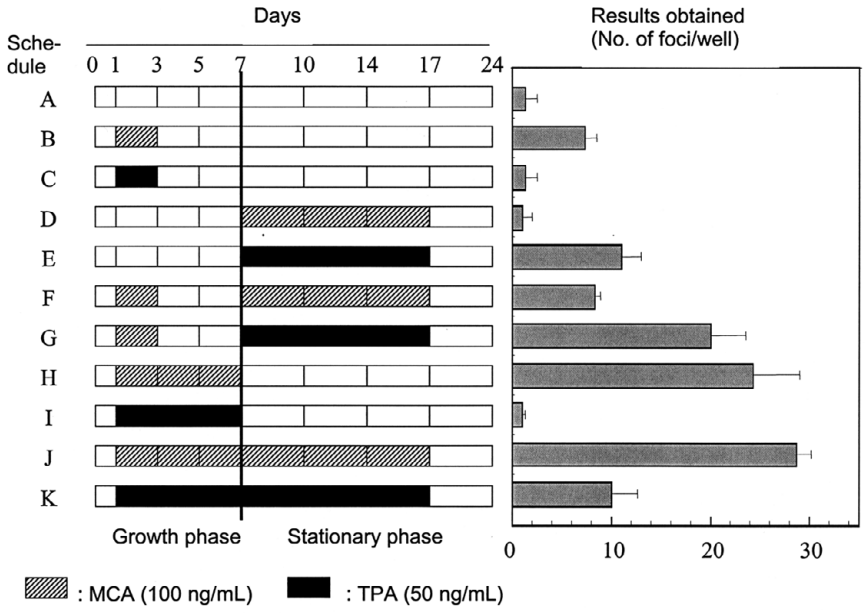


Fig. 4: MCA and/or TPA treatment under various time schedules on Bhas 42 cells

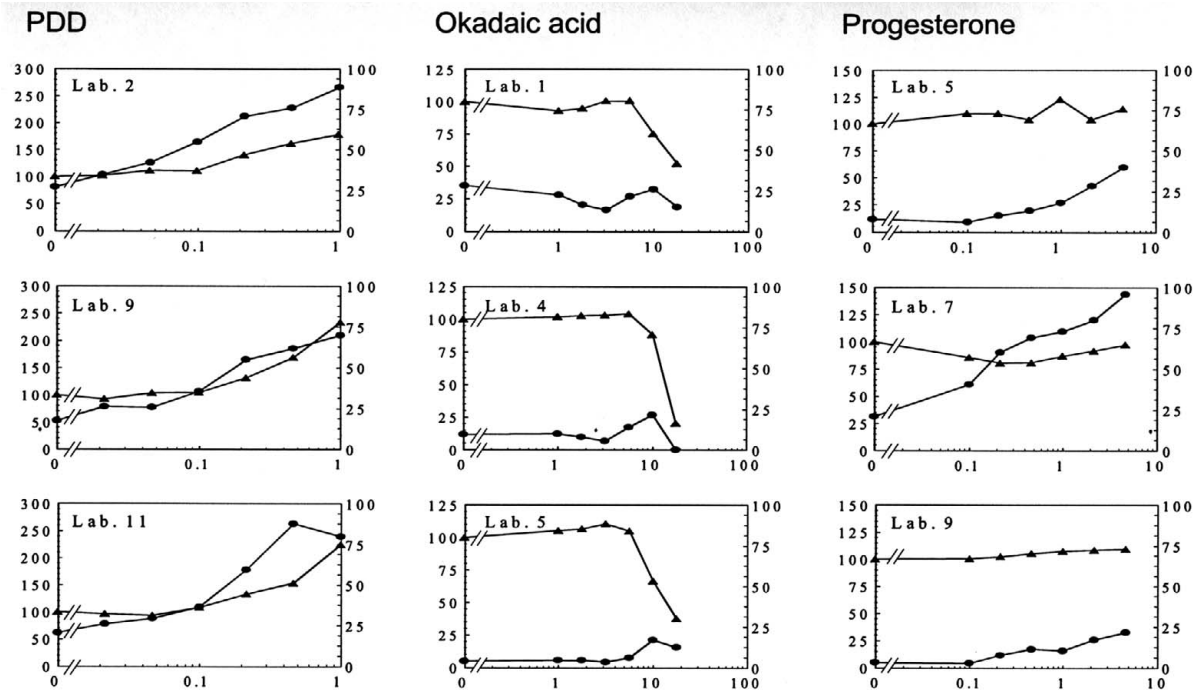


Fig. 5: Representative results in inter-laboratory collaborative study of Bhas 42 cell assay

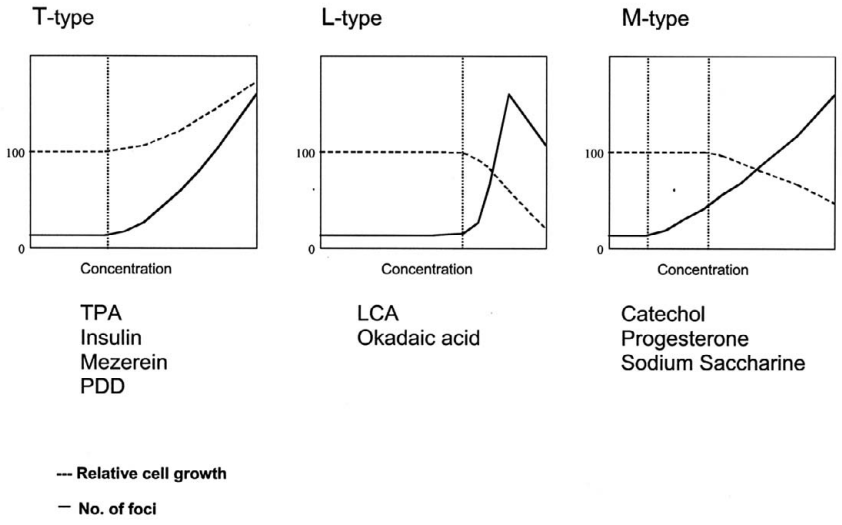


Fig. 6: Three types of chemicals showing positive response in Bhas 2nd-stage assay

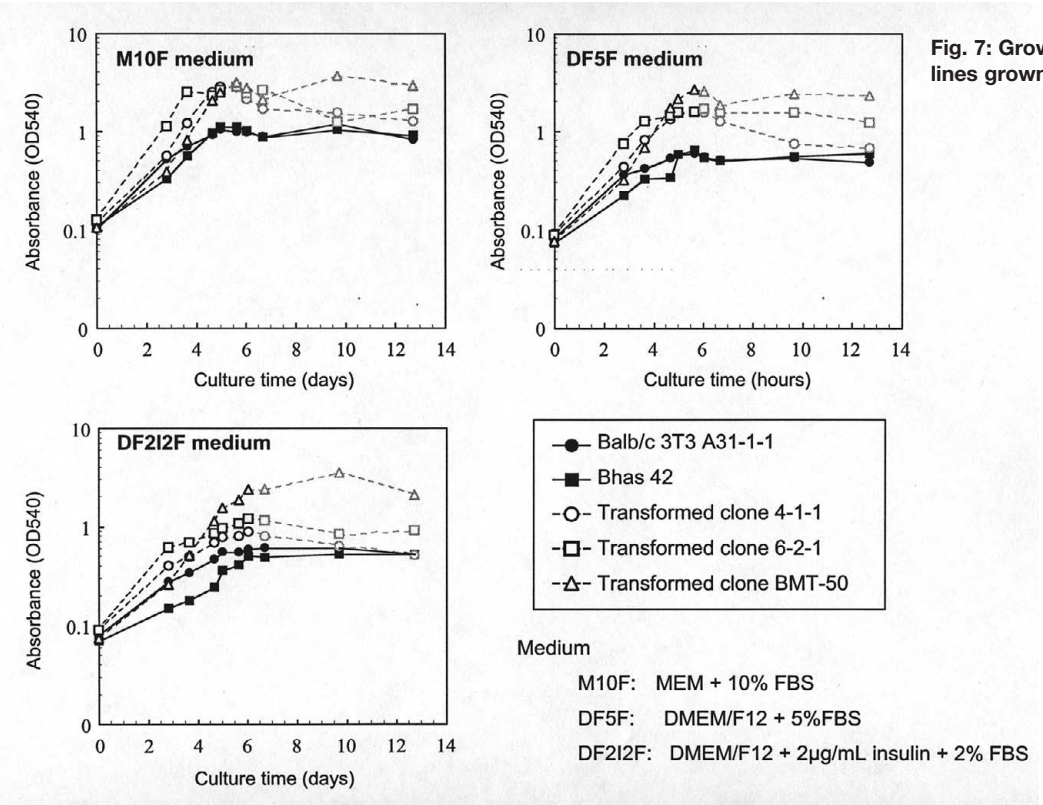


Fig. 7: Growth curves of various cell lines grown in different media

be used in the transformation assay after the preliminary cell growth assay, as well as to elucidate the mechanisms of tumour promotion.

Tests on various chemicals

Various chemicals were tested, and the results are summarised in table 3. Known tumour initiators, mitomycin C and methyl methanesulphonate, were demonstrated to be positive in the 1st-stage assay and negative in the 2nd-stage assay. Metabolites suggested to be tumour promoters were tested in the 2nd-stage assay. In relation to LCA, other tested cholic acids, cholic acid, deoxycholic acid and chenodeoxycholic acid, tested positive. Tryptophan metabolites were suggested to be tumour promoters of the bladder. 3-Hydroxykynurenine showed highly potent activity in inducing foci in the 2nd-stage assay.

Tests on non-genotoxic carcinogens revealed that 17β-estradiol was equivocal in the 1st-stage assay. Methapyrilene hydrochloride and zinc chloride were positive in the 2nd-stage assay. Diethylstil-bestrol, phenobarbital and reserpine were negative at present.

Tab. 3: Results of transformation assays on various chemicals

	1st-stage assay	2nd-stage assay
Mitomysin C	+	-
Methyl methanesulfonate	+	-
Cholic acid	Not tested	+
Deoxycholic acid	Not tested	+
Chenodeoxycholic acid	Not tested	+
Kynurenine	Not tested	+/-
3-Hydroxy kynurenine	Not tested	+
Anthranilic acid	Not tested	-
3-Hydroxy anthranilic acid	Not tested	-
Diethylstilbestrol	-	-
17β Estradiol	+/-	-
Methapyrilene hydrochloride	-	+
Phenacetin	-	-
Phenobarbital	-	-
Reserpine	-	-
Zinc chloride	-	+

The number of chemicals examined was limited, but the results demonstrated that some non-genotoxic carcinogens could be detected in the Bhas assays. Because of the simplicity of the procedure, it is anticipated that further investigations on Bhas cells will open ways of not only detecting non-genotoxic carcinogens but also elucidating mechanisms of carcinogenesis.

Growth of various cell lines in different media

Various cell lines were cultured in different media, and their growth patterns were compared (fig 7). The growth pattern of Bhas 42 cells was similar to that of BALB/c 3T3 cells in M10F and DF5F media. After both cell lines reached relatively low saturation, they entered into stationary phase, revealing non-transformed phenotype. In contrast, transformed derivatives of BALB/c 3T3 cells (4-1-1, 6-2-1 and BMT-50 clones) grew better than the non-transformed cells, reached high saturation density, and then detached from the culture surface. Thereafter, transformed cell clones resumed growth, but repeated detachment and regeneration. When saturation densities of transformed cell lines were compared to those of the non-transformed cell lines, the saturation densities differed 2- to 4-fold. The difference of this increase was greater in DF5F medium than M10F medium.

In general, growth of cell lines in DF2I2F medium was lower than in other media, and that of Bhas 42 was lowest. 2% FBS in the medium was not enough for the growth of Bhas 42 cells, even with the addition of 2 μg/ml insulin.

References

Asada, S., Sasaki, K., Tanaka, N. et al. (2005). Detection of initiating as well as promoting activity of chemicals by a novel cell transformation assay using v-Ha-ras-transfected BALB/c 3T3 cells (Bhas 42 cells). *Mutation Res. 588(1)*, 7-21.

Berenblum, I. (1941). The cocarcinogenic action of croton resin. *Cancer Res. 1*, 44-48.

IARC/NCI/EPA Working Group (1985). Cellular and molecular mechanisms of cell transformation and standardization of transformation assays of established cell lines for prediction of carcinogenic chemicals: Overview and recommended protocols. *Cancer Res. 45*, 2395-2399.

Ohmori, K., Sasaki, K., Asada, S. et al. (2004). An assay method for the prediction of tumour promoting potential of chemicals by the use of Bhas 42 cells. *Mutat. Res. 557*, 191-202.

Ohmori, K., Umeda, M., Tanaka, N. et al. (2005). An inter-laboratory collaborative study by the non-genotoxic carcinogen study group in Japan, on a cell transformation assay for tumour promoters using Bhas 42 cells. *ATLA 33*, 619-639.

Saotome, K., Morita, H. and Umeda, M. (1989). Cytotoxicity test with simplified crystal violet staining method using microtitre plates and its application to injection drugs. *Toxic. in Vitro 3*, 317-321.

Sasaki, K., Mizusawa, H. and Ishidate, M. (1988). Isolation and characterization of ras-transfected BALB/3T3 clone showing morphological transformation by 12-O-tetradecanoyl-phorbol-13-acetate. *Jpn. J. Cancer Res. 79*, 921-930.

Sasaki, K., Mizusawa, H., Ishidate, M. and Tanaka, N. (1990). Establishment of a highly reproducible transformation assay of a ras-transfected BALB 3T3 clone by treatment with promoters. *Basic Life Sci. 52*, 411-416.

Acknowledgements

This presentation consisted of results obtained in Hatano Research Institute, and in inter-laboratory collaboration with the non-genotoxic carcinogen study group. The study was supported by a grant of the Long-range Research Initiative (LRI) from the Japan Chemical Industry Association (JCIA).

Correspondence to

Makoto Umeda
729-5 Ochiai
Hadano
Kanagawa Pref. 257-0025
Japan
e-mail: GZH04502@nifty.ne.jp

A Study of Connexin Expression in Monolayer Cultures of Primary Rat Hepatocytes

Mathieu Vinken[1,3], Tamara Vanhaecke[1], Peggy Papeleu[1], Sarah Snykers[1], Tom Henkens[1], Greetje Elaut[1], James Kevin Chipman[2] and Vera Rogiers[1]

[1]Department of Toxicology, Vrije Universiteit Brussel (VUB), Brussels, Belgium, [2]School of Biosciences, University of Birmingham, Birmingham, United Kingdom, [3]Mathieu Vinken is a doctoral research fellow of the Fund for Scientific Research Flanders (FWO-Vlaanderen), Belgium

Summary

Dedifferentiation in primary cultures of adult rat hepatocytes was studied at the level of connexin expression. It was found that Cx32 protein levels gradually declined, whereas Cx26 production remained relatively stable during seven days of cultivation. Cx43, a connexin species produced by foetal hepatocytes but not by their adult counterparts, progressively appeared and was present both at the cell membrane surface and in the cytosol. Overall, this connexin switching process was associated with low gap junctional intercellular communication activity. These data provide new insights into hepatocyte dedifferentiation and aid in the optimisation of long-term primary hepatocyte cultures.

Keywords: primary hepatocyte, gap junctional intercellular communication, connexin

Introduction

Co-operation between hepatocytes is a prerequisite for normal liver functioning. In particular, direct communication via gap junctions is known to control several liver-specific functional traits including xenobiotic biotransformation (Hamilton et al., 2001; Shoda et al., 1999). Basically, this results from their capacity to mediate the exchange of small and hydrophilic molecules (e.g. Ca^{2+} and cAMP) between hepatocytes. This flux is called the gap junctional intercellular communication (GJIC). Gap junctions are composed of two connexons, in turn built up by six connexin (Cx) proteins. In liver, hepatocytes express both Cx32 and Cx26, whereas Cx43 is the only connexin species produced by non-parenchymal hepatic cells (Saez et al., 2003; Spray et al., 1994).

Primary hepatocyte cultures are valuable *in vitro* tools for pharmaco-toxicological testing of chemicals. A major shortcoming, however, includes the progressive deterioration of the hepatocyte-specific phenotype (Papeleu et al., 2002). This dedifferentiation process is also reflected at the level of GJIC. Indeed, GJIC between primary cultured hepatocytes rapidly declines, ultimately resulting in the loss of liver-specific functionality (Stoehr and Isom, 2003).

Here, we explore the mechanisms that underlie decreased GJIC between hepatocytes by studying connexin expression at the transcriptional, translational and activity level. These data are of interest for the development of new strategies to reduce dedifferentiation in primary hepatocyte cultures.

Materials and methods

Animals, hepatocyte isolation and cultivation

Procedures for the isolation and cultivation of rat hepatocytes were approved by the local ethical committee of the Vrije Universiteit Brussel. Hepatocytes (viability >80%) were isolated from adult male outbred (200-250 g) Sprague-Dawley rats (Charles River Laboratories, Belgium) by use of a two-step collagenase method (Papeleu et al., 2005), and cultivated in a monolayer configuration. The medium used was DMEM containing 0.5 IU/ml insulin, 7 ng/ml glucagon, 292 mg/ml L-glutamine, antibiotics (7.3 IU/ml benzyl penicillin, 50 µg/ml kanamycin monosulphate, 10 µg/ml sodium ampicillin, 50 µg/ml streptomycin sulphate) and 10% v/v foetal bovine serum. After 4 h, the medium was removed and renewed with fresh medium supplemented with 25 µg/ml hydrocortisone hemisuccinate and 0.25 µg/ml amphotericin B. After 24 h, the medium was renewed daily with serum-free medium. Samples were taken on days 1, 4 and 7 of the cultivation time.

RNA extraction and RT-PCR

Cells were harvested from culture plates by scraping and washed twice with cold PBS. Total cellular RNA was extracted using the SV Total RNA isolation system (Promega, USA) according to the manufacturer's instructions, and samples were subsequently subjected to DNase treatment (Ambion, USA). Reverse transcription of approximately 2 µg of total RNA was carried out using the iScript cDNA synthesis kit (Bio-Rad, Germany). For

PCR, 2 µg cDNA was amplified by means of the iTaq DNA polymerase kit (Bio-Rad, Germany). PCR primers used are listed in table 1. PCR conditions included, for Cx26: 40 cycles (1 min at 95°C, 1 min at 53.2°C, 1 min at 72°C); for Cx32: 35 cycles (1 min at 95°C, 1 min at 59.1°C, 1 min at 72°C); for Cx43: 40 cycles (1 min at 95°C, 1 min at 59.6°C, 1 min at 72°C); for GAPDH: 25 cycles (1 min at 95°C, 1 min at 52.7°C, 1 min at 72°C), each followed by final elongation at 72°C for 7 min. Samples were separated by electrophoresis on a 1.8% w/v agarose gel. Visualisation of the products was performed with ethidium bromide (0.005% v/v) under UV transillumination.

Preparation of cell lysates and immunoblotting

Cells were harvested from culture plates by scraping, washed with cold PBS, and lysed in modified Tween-20 buffer (Albrecht and Hansen, 1999). Protein concentrations were determined according to the Bradford procedure (Bradford, 1976), using a commercial kit (Bio-Rad, Germany) with bovine serum albumin as standard. Proteins (25 or 50 µg) were resolved on SDS-PAGE (7.5%, 10% or 12%) and blotted afterwards onto nitrocellulose membranes (Amersham Pharmacia Biotech, UK). Following blocking with fat-free milk, membranes were incubated with appropriate concentrations of primary antibodies (tab. 2) and HRP-labelled secondary antibodies (Dakocytomation, Denmark). Proteins were detected using the ECL detection system (Amersham Pharmacia Biotech, UK). Specificity of the primary antibodies was tested by means of positive controls, i.e. rat exocrine pancreatic tissue (10 µg protein) for Cx26 and Cx32, and rat heart homogenate (100 µg protein) for Cx43. Immunoreactivity of the Cx43 antibody in primary rat hepatocytes was further evaluated by incubating 7 day-old cell lysates overnight at 37°C with 30 IU alkaline phosphatase (AP) from calf intestine (Roche, Germany).

Tab. 1: Primers used for RT-PCR analysis. (F: forward primer; R: reverse primer)

Target	Primer sequence	Product size
Cx26	F: 5'-CGCGGCCGTCCGCTCTCCCAA-3' R : 5'-GAAGTAGTGGTCGTAGCACAC-3'	481 bp
Cx32	F : 5'-AGGTGTCGCAGTGCCAGGGAG-3' R : 5'-CCCGTGCCCCTCAAGCCGTAG-3'	373 bp
Cx43	F : 5'-GCGTGAGGAAAGTACCAAAC-3' R : 5'-GTGAAGCCGCCGCCAAAGTTG-3'	128 bp
GAPDH	F : 5'-ACCACAGTCCATGCCATCAC-3' R: 5'-TCCACCACCCTGTTGCTGTA-3'	452 bp

Immunocytochemistry

Hepatocytes were fixed with cold ethanol, permeabilised with Triton X-100, and blocked with fat-free milk. Following incubation with appropriate concentrations of primary antibodies (tab. 2) and FITC-labelled or TRITC-conjugated secondary antibodies (Zymed, USA), samples were mounted with DAPI-containing Vectashield (Vector Laboratories, USA). Detection of proteins was performed by fluorescence microscopy (Leica DMR/XA).

GJIC assay

GJIC was monitored by means of the scrape loading/dye transfer assay (el-Fouly et al., 1987) using 0.1% w/v Lucifer Yellow CH (Sigma, USA). Dye transfer was evaluated by fluorescence microscopy (Leica DMR/XA).

Results

Specificity of the primary antibodies

Connexin nomenclature is based upon molecular weight, as predicted by cDNA sequencing (Saez et al., 2003). The Cx26 antibody indeed recognised a 26 kDa protein in rat hepatocytes and in rat exocrine pancreatic cells (fig. 1A). In both cell types, however, the Cx32 antibody detected a protein of approximately 27 kDa (fig. 1A). Similar results have also been obtained by others and this finding might be explained by the occurrence of partial proteolysis of the native Cx32 protein during protein extraction, yielding a smaller (i.e. 27 kDa) protein (Kumar and Gilula, 1986; Paul, 1986).

Unlike for most other connexins, phosphorylation of Cx43 can be detected by SDS-PAGE analysis, appearing as a multiple-band signal (VanSlyke and Musil, 2000). The Cx43 antibody used here was claimed to specifically recognise the unphosphorylated Cx43 form, located at 41 kDa. When applied to rat heart tissue, a natural source of Cx43, the antibody indeed displayed single band immunoreactivity as shown in figure 1B and by others (Nagy et al., 1997). In 7 day-old primary rat hepatocytes, however, this antibody detected two closely related proteins at 41 kDa and 43 kDa (fig. 1B). Therefore, day 7 samples were treated overnight with AP. Upon subsequent immunoblot analysis, the upper band (43 kDa) became less intense (fig. 1B), suggesting that this signal corresponds with a phosphorylated Cx43 form. This finding is in agreement with the experiments of Cruciani and Mikalsen. In their model, a fibroblast cell line system, this antibody detected several phosphorylated Cx43 forms (Cruciani and Mikalsen, 1999). Thus, the 13-8300 antibody is

Tab. 2: Primary antibodies used for immunocytochemistry and/or immunoblotting.

Antigen	Supplier	Specification	Species	Type	Dilution
Cx26	Zymed	71-0500	rabbit	polyclonal	1/250
Cx32	Sigma	C3470	rabbit	polyclonal	1/500
Cx43	Zymed	13-8300	mouse	monoclonal	1/100
GAPDH	Abcam	ab8245	mouse	monoclonal	1/20000

able to recognise at least one phosphorylated Cx43 form, next to its unphosphorylated counterpart, depending on the experimental model used.

Expression of connexins in primary cultures of rat hepatocytes

Connexin expression in primary cultured rat hepatocytes was studied both at the translational and the transcriptional level. In immunoblot experiments, rat liver homogenate (LH) was included in order to have reference data for connexin expression in hepatic tissue *in vivo*.

As shown in figure 2, Cx32 protein levels decreased in primary cultures of rat hepatocytes. Cx26 protein expression, however, remained relatively stable during the whole cultivation period. In contrast, Cx43, which is not expressed by adult hepatocytes *in vivo*, progressively appeared in primary cultured rat hepatocytes. These alterations in connexin expression were also observed at the mRNA level, albeit to a lesser extent (fig. 3).

Localisation of connexins in primary cultures of rat hepatocytes

Immunocytochemistry was used to study connexin localisation during cultivation of hepatocytes (fig. 4). On day 1, Cx32 and Cx26 were mainly located in the cytosol of hepatocytes. Most likely, this is an inevitable consequence of the isolation procedure. Membrane localisation of both connexins was restored on day 4. Weak signals were observed on day 7, especially in case of Cx32. In contrast, Cx43 staining became evident on day 4, and was clearly present on day 7. Cx43 immunoreactivity was observed both at the cell membrane surface and in the cytosol of hepatocytes (fig. 4).

GJIC in primary cultures of rat hepatocytes

GJIC was evaluated by using the scrape loading/dye transfer method (fig. 5). On day 1, dye transfer between hepatocytes was limited. This finding corresponds with the cytosolic localisation of connexins observed at that time point. Dye transfer was slightly increased on day 4, but remained limited on day 7 (fig. 5).

Fig. 1 A

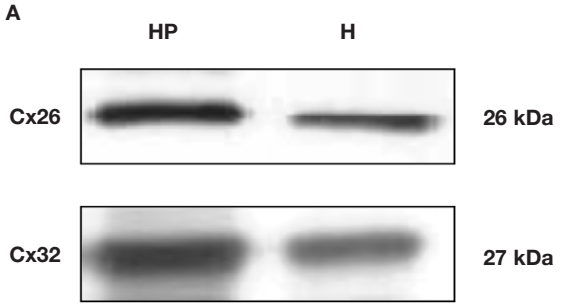


Fig. 1 B

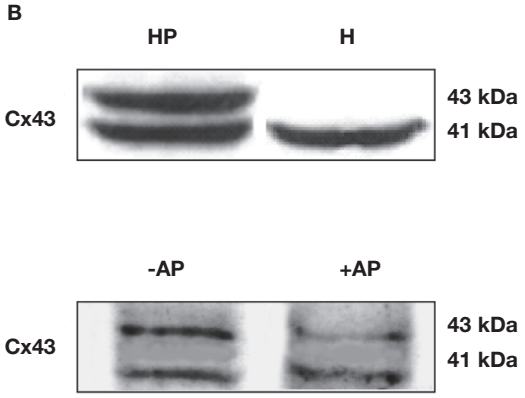


Fig. 1 A and B: Specificity of the primary antibodies.
(A) Immunoreactivity of the 71-0500 (upper panel) and C3470 (lower panel) antibodies in 1 day-old primary rat hepatocytes (HP) and in rat exocrine pancreatic cells (P). (B) Upper panel: immunoreactivity of the 13-8300 antibody in 7 day-old primary rat hepatocytes (HP) and in rat heart homogenate (H). Lower panel: immunoreactivity of the 13-8300 antibody in 7 day-old primary rat hepatocytes after overnight incubation at 37°C in the absence (-) or presence (+) of 30 IU alkaline phosphatase (AP).

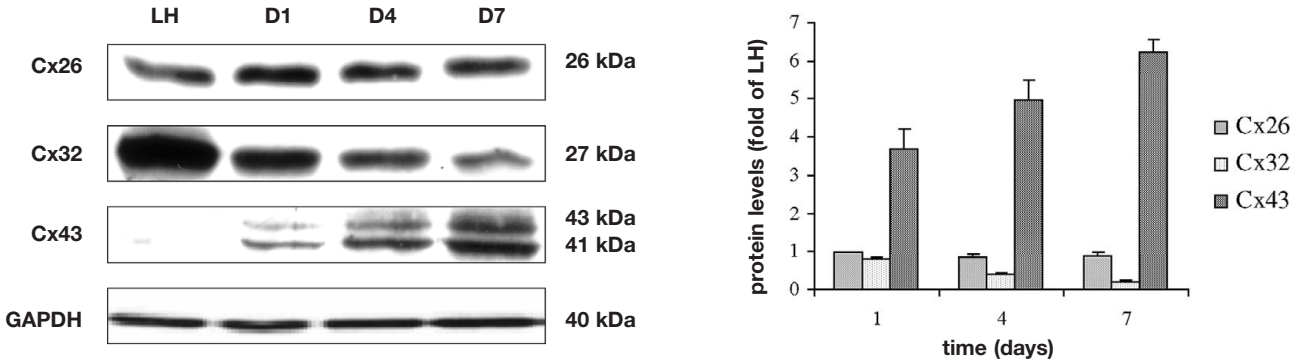


Fig. 2: Connexin protein levels in primary cultures of rat hepatocytes.
Samples were taken on days 1, 4 and 7 (D1, D4, D7) of the cultivation time and subjected to immunoblot analysis. Semi-quantification of immunoblot analyses was performed by measuring band intensities using laser densitometry. Results are expressed as fold of liver homogenate (LH) and are mean values ± standard deviation of three independent experiments.

Conclusions and perspectives

In this study, it was investigated whether the deterioration of GJIC in primary cultures of hepatocytes, known to underlie dedifferentiation, is associated with alterations in connexin expression. It was found that Cx32 protein levels rapidly decline, whereas the Cx26 production remained relatively stable during seven days of cultivation. In agreement with previous reports (Stutenkemper et al., 1992; Willecke and Haubrich, 1996), we showed that Cx43, a connexin species produced by foetal hepatocytes but not by their adult counterparts, progressively appeared in culture. Cx43 was present both at the cell membrane and in the cytosol of hepatocytes, thereby questioning its functional relevance. In fact, the molecular basis of Cx43 re-expression in primary cultured hepatocytes remains obscure. Possibly, this results from altered *cis/trans* regulation of its gene expression. Upon isolation of hepatocytes, *c*-fos and *c*-jun productions are induced (Etienne et al., 1988; Loyer et al., 1996). These proto-oncogenes dimerise to form the transcription factor activator protein-1 (AP-1), which is known to control Cx43 expression (Echetebu et al., 1999). In rat myometrium, AP-1 has been shown to induce Cx43 expression under stress conditions (Lefebvre et al., 1995). Although no solid scientific data are presently available, this scenario might also hold for isolated primary hepatocytes in culture.

Several groups have focused on the re-establishment of GJIC in primary hepatocyte cultures. Three strategies are currently followed to restore hepatocyte GJIC *in vitro*, namely (i) the introduction of an extracellular matrix, by cultivating hepatocytes on one layer or between two layers of extracellular matrix proteins (Spray et al., 1987), (ii) the establishment of cell-cell contacts, by co-cultivating hepatocytes with another cell type (Mesnil et al., 1993), and (iii) the addition of GJIC-promoting molecules to the medium of cultured hepatocytes (Yoshizawa et al., 1997).

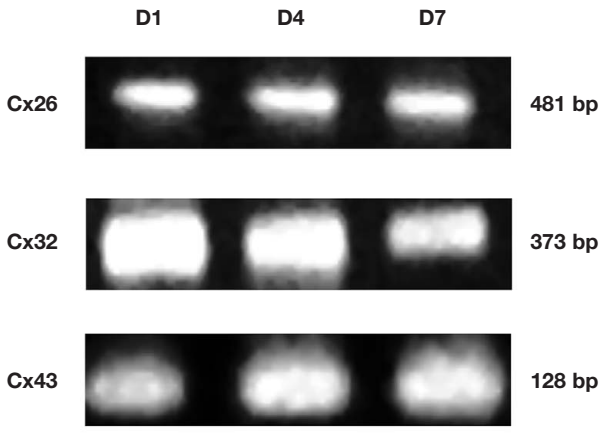


Fig. 3: Connexin gene expression in primary cultures of rat hepatocytes.
Samples were taken on days 1, 4 and 7 (D1, D4, D7) of the cultivation time and subjected to RT-PCR analysis as described in "Materials and methods".

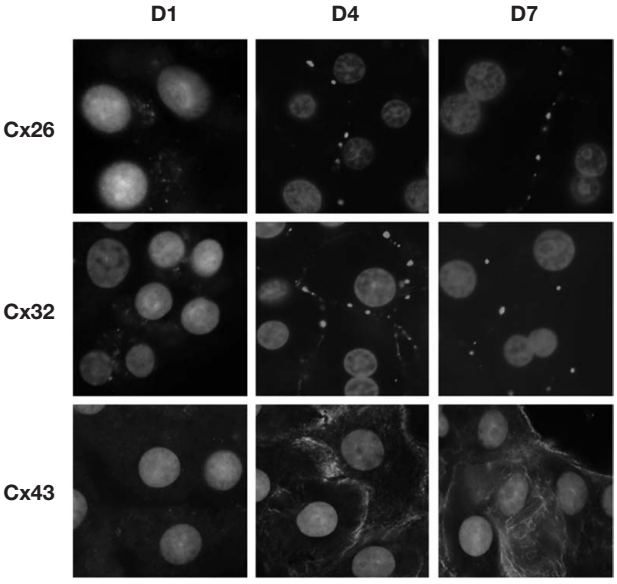


Fig. 4: Connexin protein localisation in primary cultures of rat hepatocytes.
Samples were taken on days 1, 4 and 7 (D1, D4, D7) of the cultivation time and subjected to immunocytochemistry analysis.

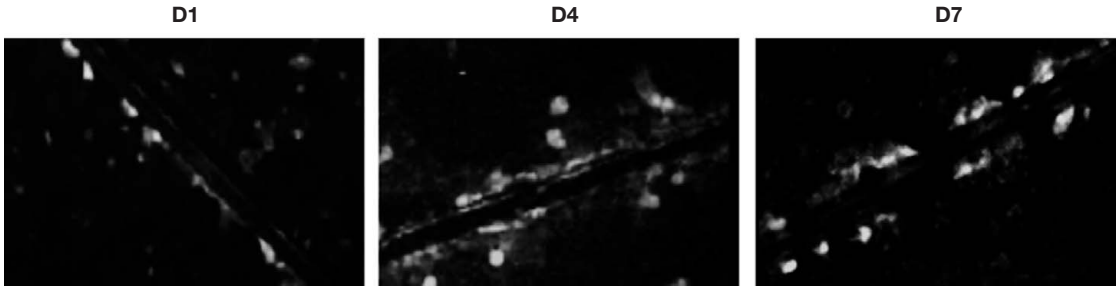


Fig. 5: GJIC in primary cultures of rat hepatocytes.
Samples were taken on days 1, 4 and 7 (D1, D4, D7) of the cultivation time and subjected to scrape loading/dye transfer analysis.

Our group is currently exploring the use of histone deacetylase inhibitors as medium supplements for primary hepatocytes. With respect to gap junctions, we have shown that trichostatin A, a prototype histone deacetylase inhibitor, promotes Cx32 expression and, consequently, GJIC (Vinken et al., in press). Further efforts to regain GJIC and thus to reduce dedifferentiation in primary cultures of hepatocytes should be strongly encouraged, as they contribute to the optimisation of liver-based *in vitro* models for long-term pharmaco-toxicological testing.

References

Albrecht, J. and Hansen, L. (1999). Cyclin D1 promotes mitogen-independent cell cycle progression in hepatocytes. *Cell Growth Differ. 10*, 397-404.

Bradford, M. M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem. 72*, 248-254.

Cruciani, V. and Mikalsen, S. V. (1999). Stimulated phosphorylation of intracellular connexin 43. *Exp. Cell Res. 251*, 285-298.

Echetebu, C. O., Ali, M., Izban, M. G. et al. (1999). Localization of regulatory protein binding sites in the proximal region of human myometrial connexin 43 gene. *Mol. Hum. Reprod. 5*, 757-766.

El-Fouly, M. H., Trosko, J. E. and Chang, C. C. (1987). Scrape-loading and dye transfer. A rapid and simple technique to study gap junctional intercellular communication. *Exp. Cell Res. 168*, 422-430.

Etienne, P. L., Baffet, G., Desvergne, B. et al. (1988). Transient expression of c-fos and constant expression of c-myc in freshly isolated and cultured normal adult rat hepatocytes. *Oncogene Res. 3*, 255-262.

Hamilton, G. A., Jolley, S. L., Gilbert, D. et al. (2001). Regulation of cell morphology and cytochrome P450 expression in human hepatocytes by extracellular matrix and cell-cell interactions. *Cell Tissue Res. 306*, 85-99.

Kumar, N. M. and Gilular, N. B. (1986). Cloning and characterization of human and rat liver cDNAs coding for a gap junction protein. *J. Cell Biol. 103*, 767-776.

Lefebvre, D. L., Piersanti, M., Bai, X. H. et al. (1995). Myometrial transcriptional regulation of the gap junction gene, connexin-43. *Reprod. Fertil. Dev. 7*, 603-611.

Loyer, P., Cariou, S., Glaise, D. et al. (1996). Growth factor dependence of progression through G1 and S phases of adult rat hepatocytes in vitro: evidence of a mitogen-dependent restriction point in mid-late G1. *J. Biol. Chem. 271*, 11484-11492.

Mesnil, M., Piccoli, C. and Yamasaki, H. (1993). An improved long-term culture of rat hepatocytes to detect liver tumour-promoting agents: results with phenobarbital. *Eur. J. Pharmacol. 248*, 59-66.

Nagy, J. I., Li, WE. I., Roy, C. et al.(1997). Selective monoclonal antibody recognition and cellular localization of an unphosphorylated form of connexin43. *Exp. Cell Res. 236*, 127-136.

Papeleu, P., Elaut, G., Rogiers, V. and Vanhaecke, T. (2002). Cell cultures as in vitro tools for biotransformation studies. In S. G. Pandalai (ed.), *Recent Research Developments in Drug Metabolism and Disposition vol. 1* (199-234). Kerala: Transworld Research Networks.

Papeleu, P., Vanhaecke, T., Henkens, T. et al. (2005). Isolation of rat hepatocytes. In I. R. Phillips and E. A. Shephard (eds.), *Methods in Molecular Biology* (vol. 320): Cytochrome P450 protocols 2nd ed. (229-237). Totowa: Humana Press.

Paul, D. L. (1986). Molecular cloning of cDNA for rat liver gap junction protein. *J. Cell Biol. 103*, 123-134.

Saez, J. C., Berthoud, V. M., Branes, M. C. et al. (2003). Plasma membrane channels formed by connexins: their regulation and functions. *Physiol. Rev. 83*, 1359-1400.

Shoda, T., Mitsumori, K., Onodera, H. et al. (1999). The relationship between decrease in Cx32 and induction of P450 isozymes in the early phase of clofibrate hepatocarcinogenesis in the rat. *Arch. Toxicol. 73*, 373-380.

Spray, D. C., Fujita, M., Saez, J. C. et al. (1987). Proteoglycans and glycosaminoglycans induce gap junction synthesis and function in primary liver cultures. *J. Cell Biol. 105*, 541-551.

Spray, D. C., Saez, J. C., Hertzberg, E. L. and Dermietzel, R. (1994). Gap junctions in liver: composition, function, and regulation. In I. M. Arias, I. L. Boyer, N. Fausto, W. B. et al. (eds.), T*he liver: Biology and Pathobiology 2nd ed.* (951-967). New York: Raven Press Ltd.

Stoehr, S. A. and Isom, H. C. (2003). Gap junction-mediated intercellular communication in a long-term primary mouse hepatocyte culture system. *Hepatology 38,* 1125-1135.

Stutenkemper, R., Geisse, S., Schwarz, H. J. et al. (1992). The hepatocyte-specific phenotype of murine liver cells correlates with high expression of connexin32 and connexin26 but very low expression of connexin43. *Exp. Cell Res. 201*, 43-54.

VanSlyke, J. . and Musil, L. S. (2000). Analysis of connexin intracellular transport and assembly. *Methods 20*, 156-164.

Vinken, M., Henkens, T., Vanhaecke, T. et al. (2006). Trichostatin A triggers gap junctional intercellular communication in primary cultures of adult rat hepatocytes. *Toxicol. Sci*. in press.

Willecke, K. and Haubrich, S. (1996). Connexin expression systems: to what extent do they reflect the situation in the animal? *J. Bioenerg. Biomembr. 28*, 319-326.

Yoshizawa, T., Watanabe, S., Hirose, M. et al. (1997). Dimethylsulfoxide maintains intercellular communication by preserving the gap junctional protein connexin 32 in primary cultured hepatocyte doublets from rats. *J. Gastroenterol. Hepatol. 12*, 325-330.

Acknowledgements

This work was supported by grants from the Fund of Scientific Research Flanders (FWO-Vlaanderen), Belgium, the Research Council of the Vrije Universiteit Brussel (OZR-VUB), Belgium, and the European Union (project number 504761).

Correspondence to

Mathieu Vinken
Vrije Universiteit Brussel (VUB)
Department of Toxicology
Laarbeeklaan 103
1090 Brussels, Belgium
e-mail: mvinken@vub.ac.be

Authors

Keywords

ALTEX
www.altex.ch

A quarterly journal for new paths in biomedical science
Official organ of the Middle European Society for Alternative Methods to Animal Testing (MEGAT)

Bd. 23 (Special Issue 2006)

Editor:
Society ALTEX Edition, Zurich, Switzerland
President: Nina Schweigert (Zurich),
Vice president: Helmut Appl (Wien),
Society Board: Thomas Hartung (Konstanz).
Secretary and Editor in Chief: Franz P. Gruber.

Editorial Office:
Franz P. Gruber, Horst Spielmann, Susanne Scheiwiller, Irène Hagmann (Zurich) and Sonja von Aulock (Konstanz).
ALTEX, Hegarstr. 9, P.O.B. 1766, 8032 Zurich, Switzerland
Phone: +41-44-380 08 30,
Fax: +41-44-422 80 10,
E-mail: info@altex.ch.

Advisory Board:
Peter Bossard (Horw, Switzerland)
Miroslav Cervinka (Hradec Cralove, Czech Republic)
Nicole Clemann (Basel, Switzerland)
Klaus Cußler (Langen, Germany)
Andrzej Elzanowski (Wroclaw, Poland)
Friedrich Harrer (Salzburg, Austria)
Thomas Hartung (Konstanz, Germany)
Jan G. Hengstler (Leipzig, Germany)
Jane Huggins (Plainsboro, USA)
Coenraad F. M. Hendriksen (Bilthoven, The Netherlands)
Claudia Mertens (Winterthur, Switzerland)
Brigitte Rusche (Neubiberg, Germany)
Walter Pfaller (Innsbruck, Austria)
Harald Schöffl (Linz, Austria)
Gotthard M. Teutsch (Bayreuth, Germany)
Jan van der Valk (Utrecht, The Netherlands)
Angelo Vedani (Basel, Switzerland)

ALTEX Online:
http://www.altex.ch
English summaries:
http://altweb.jhsph.edu/publications/journals/altex/altex-index.htm,
Webmaster: caat@jhsph.edu
ALTEX is indexed in MEDLINE, Current Contents®, SciSearch® and ISI Document Solution® and EMBASE.

ALTEX is published with the support of the society members and the following sponsors:
Foundation for animal-free Research, FFVFF, Zurich, Switzerland (www.ffvff.ch); Doerenkamp-Zbinden Foundation, Zurich, Switzerland (www.doerenkamp-zbinden.org); MEGAT (Middle European Society for Alternative Methods to Animal Testing), Linz/Wien, Austria (www.zet.or.at/MEGAT/index.htm); Tierschutz Bund Zürich, Dübendorf, Switzerland (www.tierschutzbund-zuerich.ch); Zürcher Tierschutz, Zurich, Switzerland (www.zuerchertierschutz.ch); Foundation SET, Mainz, Germany (www.tierversuche-ersatz.de); German Animal Welfare Federation, Bonn, Germany (www.tierschutzbund.de); Menschen für Tierrechte – Tierversuchsgegner Baden-Württemberg e.V., Stuttgart, Germany (www.tierrechte-bw.org); Ligue Suisse Contre la Vivisection, Thonex, Switzerland (www.lscv.ch); Ärztinnen und Ärzte für Tierschutz in der Medizin, Zurich, Switzerland (www.aerztefuertierschutz.ch); Foundation Biografics Laboratory 3R, Basel, Switzerland (www.biograf.ch); Fondation Egon Naef, Thonex, Switzerland (www.fondation-naef.com).

Layout: H. P. Hösli, Zurich, Switzerland
Print: Stürtz GmbH, Würzburg, Germany

Distribution:
ALTEX, Schützenstr. 14, 78462 Konstanz, Germany;
ALTEX, Postfach 100125, 78401 Konstanz, Germany,
Phone: +49-7531 16614,
Mobile: +49-1726272224,
Fax +49-7531 25833,
E-mail: abo@altex.ch.

Publisher: Elsevier GmbH, Spektrum Akademischer Verlag, Slevogtstrasse 3-5, 69115 Heidelberg, Germany,
Phone: +49-6221-9126 324,
Fax: +49-6221-9126 370,
www.elsevier.de.

Edition: 1300

Circulation: 4 times per annum

Subscription rates:
Europe
Private subscribers 75 €
Libraries, commercial institutions 150 €
Outside Europe
Private subscribers 91 €
Libraries, commercial institutions 166 €

Subscription rates include costs of shipment (four issues/year) and taxes. Animal welfare organisations, students and subscribers from outside Europe and former Soviet Union may ask for special subscription rates.

Aims and scope

ALTEX, edited by the Swiss Society ALTEX Edition is the official journal of MEGAT, the Middle European Society for Alternative Methods to Animal Testing. ALTEX is devoted to the publication of research on the development and promotion of alternatives to animal experiments according to the 3R concept of Russell and Burch: Replace, Reduce, and Refine. ALTEX is publishing original articles, short communications, reviews as well as news and comments, meeting reports and book reviews in English and German. Animal experiments are defined by the editors as all experimental procedures using animals which may cause pain, suffering, and emotional harm to animals and which are conducted in research and education or to obtain tissues, organs, and other animal derived products.
Besides covering the biomedical aspects of animal experimentation, ALTEX is also devoted to the bioethics of the complex relationship between man and animals. Articles published in ALTEX should express a basic concern about the dignity of living creatures. ALTEX is not only aimed at developing a new approach to recognise animals as partners but it also intends to introduce a scientific sight in the discussions on animal experiments. Articles devoted to the social and ethical aspects of this topic will, therefore, be judged according to stringent scientific standards. Manuscripts submitted to ALTEX are evaluated by two reviewers. The evaluation takes into account the scientific merit of a manuscript and the contribution to animal welfare and the 3R-principle.

© Copyright: Elservier GmbH, Spektrum Akademischer Verlag, Heidelberg, Germany