



5 Discussion on best practices in AOP development

Led by Carl Westmoreland, the discussion following the first sessions focused on “best practices” for AOP development. First, consensus was reached on the basic elements needed for pathway development, including an established MIE and an AO that are linked by any number of intermediate events. One issue on which consensus was not reached was the limitations that this structure might impose on AOPs, i.e., must every pathway end with an AO at the individual or population level and must every pathway be strictly linear without branching? Biologically a chemical can result in more than one AO from a single MIE or may have multiple MIEs; however, in practice, an AOP built using specific chemicals as case examples is likely to be initially built from one MIE to a single AO, and it is envisioned that multiple, linear AOPs would likely be built rather than a complex, branched AOP. On the other hand, a more holistic view of the branch points would likely eventually be created. An additional consideration might be to group pathways with similar anchors together. Regardless of the strategy for development, an initial goal is to develop a few specific examples in such a way as to make the process generalizable (to other chemicals, to other tissues) and the resulting pathways testable.

Comparison was made between AOP development and similar pathway construction efforts following alternate frameworks, such as the biological anchoring of QSARs, MoA Human Relevance Frameworks, and toxicity pathway identification. One issue raised was whether every conceivable MIE needs a pathway, or, alternatively, does every pathway need an MIE? Thorough description of every conceivable pathway may not be necessary to capture the full spectrum of biological consequences of chemical exposure.

One distinguishing feature of AOPs from other pathway frameworks may be that the intended use of the AOP can result in different levels of required detail. As such, use should be considered a necessary and fundamental element in the development of an AOP. Since the pathways are designed for a purpose, the design will be different if you are accurately predicting toxic effects or predicting regions of safety (i.e., non-toxic exposure levels). In risk assessment for complex endpoints, it is critical to take homeostasis/recovery/repair (feedback loops) into account in the pathway. It is also critical to add quantitative information to the relationships between events – for example, with mitochondrial toxicity/ATP production, the threshold is high but the dose-response is steep. Furthermore, it may be necessary (in human health risk assessment) to take inter-individual variability into account. There may also be different classes of pathways, such as receptor-mediated (recognition) and reactivity (protein or DNA damage) pathways. Nonetheless, are these classes qualitatively different? Both quickly cause changes in

gene expression, so the pathways may actually be similar in dynamics and structure. Overall, the importance of the hypothesis or hypotheses should be emphasized; it is important to be objective and not be too attached to a particular hypothesis that might interfere with the recognition of other factors and relationships.

Depending on the intended use, the various elements of an AOP might need different levels of scrutiny. For example, a thorough description of the MIE is necessary for QSAR modeling, while a more general understanding of the intermediate events without a detailed MIE might be useful for designing integrated testing strategies. In some cases, identifying and measuring intermediate events in a pathway may be more useful, allowing the capture of multiple MIEs and broader coverage of potential routes to the AO.

A main objective of an AOP is, in part, to assist in test design, so an AOP should provide a clear organizational framework (e.g., in an integrated testing strategy) to aid the effective and accurate prediction of toxicity, and perhaps more critically, regions of safety. As part of this focus on chemical testing strategy, the OECD has developed guidance, in part based on suggestions from the PoT and JRC workshops, to provide an initial evaluation and assessment of AOPs. Participants were generally in agreement that the suggested OECD principles of data and causal relationship evaluation are appropriate. Furthermore, there was consensus that if an AOP relies entirely on comparisons with animal results (where concordance is not all that good), the system is limited. Thus, there is a desire to incorporate measures of human relevance of the AOP where possible.

Participants agreed that future activities in AOP development should include further definition and identification of intermediate events, and that establishing a library of “key” events would be useful. Whether or not an event is “key” may, of course, depend on the pathway and the data to support causality between events. In addition to a library of key events, a centralized reference or guidance and published examples of pathways are needed.⁷ This could be achieved through crowd-sourcing or gaming-specific problems and data challenges; these techniques have been shown to be good ways to creatively address complex problems in other scientific disciplines.⁸

6 Pathways relating to liver toxicity

Because liver toxicity is prevalent in acute, chronic and subchronic chemical exposure – whether due to repeated exposure to environmental chemicals or to drugs (where liver toxicity is a leading cause of removal from the market) – it is an especially appropriate subject for pathway case studies. Drug-induced liver injury (DILI) is rare, making it hard to detect in preclinical testing and clinical trials, but it can be life-threatening (requiring liver transplantation) or fatal for an affected

⁷ An AOP Knowledge Base (AOP KB) is currently being implemented as an OECD project to facilitate the contribution to and evaluation of AOP information collected by a wide range of experts. Currently a beta version is available only for OECD AOP development teams; EPA intends to release a public version in September 2014.

⁸ For example, “Fold-it” (<http://fold.it>) invites players to design or predict protein structures (Cooper et al., 2010. “Ete RNA” (<http://eterna.cmu.edu/web/>) players design RNA molecules and compete for the chance to have their molecule synthesized (Lee et al., 2013).



individual. In addition, any indication of liver toxicity during drug development, whether in animal studies or *in vitro*, may result in abandonment of the drug, and promising therapeutics are likely to be inappropriately discarded.

DILI can also be highly dependent on the patient's own genetic makeup and physiological condition. While there are described diagnostic markers for DILI, these are not specific and, more importantly, are not predictive of future occurrence. Understanding liver toxicity therefore requires an in-depth understanding of normal liver function as well as the pathway perturbations. Pathway-based approaches to liver toxicity and DILI are important for identifying predictive biomarkers.

6.1 Drug-induced liver injury: Mechanisms and types – Mathieu Vinken

While drug induced liver injury (DILI) is infrequent (1 in 10,000 to 100,000 patients per year), it is responsible for 50% of the cases of liver failure, and most of these are due to reactions to acetaminophen. The types of drugs most commonly associated with liver toxicity are antibiotics, central nervous system drugs, and non-steroidal anti-inflammatory drugs. Animal toxicity studies pick up about 50-70% of the drugs that will cause liver toxicity; *in vitro* studies will detect about 50-60%. To detect liver injury in clinical trials, the trial size would need to be in the order of 30,000 patients, rather than the standard 2,000-5,000.

DILI can be acute or chronic and be of two types: intrinsic (also called A-type; associated with the drug's activity, dose-dependent and has a latency of only hours or days), or idiosyncratic (also called B-type; not dose-dependent, has a latency of weeks to months). Idiosyncratic DILI can be further subdivided into two types, metabolic and immunologic, based on the origin of the cause. These types are often not easily distinguished, the latency period making identification of the cause difficult. Regardless of the type of DILI, there is thought to be a 3-step process, beginning with an initial cellular injury (by a parent drug and/or its metabolites) resulting in direct cell stress, induction of an immune response or direct mitochondrial inhibition; followed by mitochondrial permeability; and resulting in cell death (apoptosis or necrosis) (Russmann et al., 2009).

Acute DILI can be subdivided into cytotoxic (resulting in cellular injury) or cholestatic (resulting in a block in the flow of bile). Acute DILI often manifests as hepatic necrosis, hepatocellular steatosis and/or degeneration or acute cholestasis, whereas chronic liver injury can manifest in a number of conditions including hepatitis, steatosis, fibrosis, phospholipidosis and neoplastic lesions.

Following the 3-step model of DILI, the first step, as an example, might include direct cell stress as the cause of the initial cellular injury. Direct cell stress may cause necrosis and/or degeneration of the cell. It can also cause steatosis (a build-up of fatty acids in hepatocytes). Steatosis occurs via several mechanisms: impaired beta-oxidation (e.g., amiodarone), increased *de novo* synthesis of fatty acids (e.g., valproic acid), damage to mitochondrial DNA (e.g., zidovudine), and altered secretion of very low density lipoprotein (e.g., tetracycline). A lesion that presents similarly to steatosis histologically is phospholipidosis.

Phospholipidosis is a reversible accumulation of lipids caused by drugs with a cationic amphiphilic structure. Phospholipidosis can be caused by increased phospholipid synthesis and/or decreased phospholipase activity. The clinical relevance of phospholipidosis is not clear; it could result in liver injury or be an adaptive response to liver injury. Cholestasis can also result from direct cell stress via several mechanisms, including: inhibition of Na⁺/K⁺ ATPase (e.g., chlorpromazine), inhibition of bile salt export pump (BSEP) (e.g., glibenclamide), and agonism of MRP/P-glycoprotein (e.g., estradiol 17beta-glucuronide) but can also occur indirectly via swelling of hepatocytes and blockage of bile canaliculi.

Another possible cause of initial cellular injury in the 3-step model is induction of a specific immune response against the liver, likely involving induction of both innate and adaptive immunity. Involvement of the immune response results in production of cytokines and activation of B and T cells. Inflammation generally results, although cytokines themselves can be pro- or anti-inflammatory. To cite one example, idiosyncratic or B-type DILI involves the adaptive immune system and symptoms take 1-4 weeks to develop. One proposed mechanism for this is the hapten induction hypothesis, which theorizes that drugs (e.g., halothane, diclofenac or dihydralazine) covalently interact with proteins to form haptens. These haptens are then processed by dendritic cells and fragments are presented via major histocompatibility complex (MHC) proteins to T cells that then activate the inducible immune response (activated T cells and antibody-producing B cells). An alternative mechanism hypothesis (the pharmacological interaction (pi) concept) is that some drugs (e.g., lamotrigine, carbamazepine or sulfamethoxazole) bind directly to MHC molecules and mimic T-cell antigens, thereby activating T cells directly.

Following the initial cell injury (either direct cell death, induction of an immune response or direct mitochondrial inhibition), the second step in the three-step DILI model is induction of mitochondrial dysfunction, such as by disruption of the permeability transition pore (PTP). The PTP regulates the uptake of water and release of cytochrome C. The extent of PTP opening determines the degree of ATP depletion, which in turn determines the mode of the third and final step in the three-step DILI model, cell death. A greater depletion of ATP leads to necrosis, a lesser depletion leads to apoptosis.

Whatever the cause of DILI, be it Type-A intrinsic or Type-B idiosyncratic, repeated, chronic insults that cause cycles of cell death and re-growth will result in liver fibrosis. Liver fibrosis can be a reversible process and is induced by a range of drugs including oral contraceptives and exposure to methotrexate. Key events in fibrosis include: activation of stellate cells to myofibroblast-like cells, loss of vitamin A in stellate cells, deposition of extracellular matrix proteins, Kupffer cell activation and loss of fenestrae in sinusoidal endothelial cells.

6.2 Drug induced liver injury (DILI) – John Senior

The FDA is charged by Congress to determine whether drugs for public use have (1) effectiveness, and (2) safety. It is not really a risk/benefit ratio – but more an assessment of the difference between risk and net benefit. The liver serves as the body's



chemical engineering and control center, regulating the metabolism of internal compounds and also coping with compounds coming in from the environment, such as drugs. The liver has a remarkable capacity to adapt and regenerate, which can confound predictive models.

Approved drugs are by far the most common cause of acute liver failure in the United States. Drugs are not intended to cause harm; however, several widely used drugs cause acute liver failure and liver injury and have required regulatory action. While some drugs are more likely to cause liver injury than others, it is also true that some patients are more susceptible to the same drug and doses than others.

Acetaminophen is so widely the focus of DILI research because of its very successful marketing in the U.S. since 1986, when it was positioned as a “safe” aspirin alternative after the use of aspirin in children with “flu” symptoms was associated with onset of Reye’s Syndrome. Since many people think of it as safer – and assume that if a little is good, a lot would be better – there is a tendency to take too much, leading to toxicity.

In order to predict and prevent the occurrence of DILI, we need to know both severity and causality of hepatotoxicity. Severity measures can be misleading. Most people tolerate exposure well; as severity goes up, the number of affected individuals goes down and very few experience severe reactions or death. In addition, due to the number of factors involved, causality is very difficult to establish.

The study of DILI began following the approval, in 1997, of two drugs, troglitazone (Rezulin: approved 29 Jan 1997, Parke-Davis, a new class of anti-diabetic drugs) and bromfenac, (Duract: approved 15 July 1997, Wyeth-Ayerst, a short-term analgesic). In both cases, patients almost immediately began showing severe, fatal cases of centrilobular hepatic necrosis. Duract was removed from market in June 1998; Rezulin took longer, it was removed in March 2000.

Identifying a causal relationship between drugs and DILI is difficult and requires a lot of work. “Hy’s Law” established a relationship between biomarkers and the likelihood of DILI and includes the following findings: (1) alanine transaminase or aspartate transaminase greater than 3 times the upper limit of normal, and (2) total bilirubin of greater than 2 times the upper limit of normal, and (3) not caused by pre-existing disease. These findings, combined with a history of drug administration, would indicate that a drug is a probable cause, but clinical adjudication (differential diagnosis) is also required.

To help identify DILI cases for evaluation, FDA uses “eD-ISH” (evaluation of Drug Induced Serious Hepatotoxicity), a software program that uses the human ability of pattern recognition combined with the computer’s ability to sift through data that would take a human being onerous amounts of time.

Regarding DILI risk, questions with regulatory impact include: does a drug cause clinically significant DILI in the target treatment population? What is the clinical signature of the injury? What ranges of dose and exposure are associated with increased risk? We should ask these questions at all phases of a drug’s life cycle.

Patient susceptibility factors are a huge determining feature of DILI; there are no idiosyncratic drugs, the idiosyncrasy is in

the recipient. That is why modeling does not always give us the answers. Predictive biomarkers of DILI that identify susceptible patients are needed but do not yet exist.

6.3 Predicting serious DILI – clinical & scientific challenges – Mark Avigan

Given numerous challenges, such as variability in the onset and tempo of DILI and in the drug-specific liver profiles of patients, we need to improve quantifying and managing risk. Since only a small percentage of the exposed population reacts, what we really need are biomarkers. This is the challenge: to identify biomarkers that will allow us to create a model that can be used to predict which drugs will cause DILI, which patients are susceptible, and which cases are likely to undergo clinical resolution vs. progression while still at an early stage of mild liver injury. Any model will need to account for the fact that the liver is particularly good at repairing itself.

In terms of predicting idiosyncratic hepatotoxicants, single cell cultures and animal studies have been of little help: there is only about 50% concordance (which raises another problem: if a drug shows indications of toxicity in animals, it will not be pursued, so many possible human therapeutics may be abandoned). Improvement in prediction for some, but not all, types of hepatotoxicants may be obtained using *in vitro* co-culture systems (hepatocytes with mesenchymal cells), multi-parametric (cell viability, mitochondrial functions, Nrf-2 activation, GSH adducts, etc.) and high-throughput analyses, and screens for effects on hepatocyte efflux transporters. So far, *in vitro* drug screening for hepatotoxicity has not modeled risk effects of idiosyncratic susceptibility.

Current biomarkers of DILI have important limitations: they are non-specific and cannot predict recovery versus adverse outcome. Ideally, biomarkers would have the following characteristics: measurable non-invasively; low baseline variability in tolerant populations; appear before the onset of severe hepatotoxicity; and be accurately quantifiable with a dynamic range in susceptible populations. Pathways of toxicity and injury reveal opportunities for biomarker development.

The feasibility of finding genetic biomarkers is a function of two factors: the amplitude of the effect, and the allele frequency within the affected population. Extreme phenotypes are ways to start looking at the system: for example, people with Alpers-Huttenlocher syndrome have a 1 in 3 chance of going into liver failure if given valproic acid. Identifying the predisposing factor to liver failure development in such a population might allow extrapolation to other populations. Many susceptibility markers are MHC-related. There is also a rising interest in mitochondrial toxicity as a predictor. Because of the reliability and specificity of their release, there is increasing interest in circulating RNAs as biomarkers.

From a biosystems perspective, a key point is that to go from mild perturbation all the way to liver failure is usually going to require more than one “hit;” it is going to be a combination of multiple hits. These hits will be in a combination of drug-specific and non-specific hepatotoxicity steps in pathways leading to toxicity. From the point of view of modeling, you may therefore need both drug-specific and non-specific biomarkers.



Some serious challenges to the development of predictive models include the fact that the injury occurs in a complex organ system with many interconnected components, metabolism and clearance of the drug must be taken into account, there are competing toxicity and cytoprotective pathways activated at the same time, escape from homeostasis can lead to a rapid progression (the course of the disease can change suddenly), both innate and adaptive immunity can be involved, the “butterfly effect” can occur (slight differences in initial conditions can result in vastly different results) and different critical combinations of susceptibility factors can occur in exposed populations.

In summary, absolute predictors of DILI are unlikely. Predictive modeling of serious DILI will likely be a composite of biomarker measures that include circulating mRNA/miRNA, serum/urinary metabolites & P-G /epigenetic markers, and biomarker measures would reflect early and intermediate steps leading to DILI progression or adaptation/recovery.

6.4 Liver toxicity pathways example: Acetaminophen – Hartmut Jaeschke

Acetaminophen is a classic case study of liver toxicity and has been extensively studied for many years. Acetaminophen is a very safe drug at prescribed doses, but can be fatal in overdose, so acetaminophen represents a case where using high dose exposure in animal testing is very similar to actual human exposure scenarios. Metabolic activation by P450s is a critical first step in acetaminophen’s mechanism of action, followed by protein binding/adduct formation. In the liver, roughly 50% of acetaminophen is metabolized by glucuronosyl-transferases and excreted in the bile and plasma, another 30% is metabolized by sulfo-transferases and also excreted. 5-10% is metabolized by CYP450s into the reactive intermediate N-acetyl-p-benzoquinoneimine (NAPQI), which is either deactivated by glutathione-s-transferases and excreted, or reacts with proteins to form protein adducts. During a toxic response, cell death occurs; however, protein adducts also form in people and animals under normal therapeutic conditions. It is possible that at low doses and low levels of adduct formation, elimination mechanisms (autophagy) prevent toxicity-related damage but at higher doses adducts cannot be removed fast enough. The question of whether toxicity involves a threshold or hitting a specific critical target is currently being addressed by proteomics. Although there is inflammation associated with acetaminophen-induced liver toxicity, there is no associated immune response (the protein adducts do not behave like haptens), but it is unclear as to why this is.

Protein adducts themselves are unlikely to cause cell death, so what is the mechanism? There is evidence of mitochondrial dysfunction and endoplasmic reticular stress in mice. One hypothesis is that covalent attachment of NAPQI to proteins, especially mitochondrial proteins, initiates mitochondrial dysfunction, which leads to oxidative stress and release of mitochondrial intermembrane proteins (e.g., endonuclease G), which cause nuclear DNA fragmentation. Glutathione protects against the mitochondrial oxidant stress, which can trigger the opening of the membrane permeability transition pore, leading to collapse of the membrane potential, ATP depletion and eventually necrotic

cell death. An adaptation to the mitochondrial stress is removal of damaged organelles by autophagy (mitophagy). Autophagosomes containing damaged organelles fuse with lysosomes, degrading the damaged organelles, and release the amino acids, etc., back into the cell to be used as energy material. Cellular toxicity occurs when the cells’ ability to respond to mitochondrial stress is overwhelmed. So the mechanism would be: protein adduct formation, initial mitochondrial oxidant stress followed by activation of several upstream MAP kinases resulting in activation (phosphorylation) of c-jun-N-terminal kinase (JNK). Phosphorylated JNK translocates to the mitochondria causing the amplification of the stress response, which overwhelms the impaired endogenous defense mechanisms (GSH, autophagy), leading to cell necrosis (Jaeschke et al., 2012a).

All of the work described so far was done in mice, so what is the relevance for humans? The mechanism of toxicity was investigated in HepaRG cells, which are derived from human liver tumors, and appears to have many of the same features of the response in mice, except that there is no JNK activation in HepaRG cells and no protection by JNK inhibitors. However, acetaminophen-induced cell death in primary human hepatocytes is at least in part dependent on JNK activation. The human response in overdose patients involves the formation of protein adducts, GSH depletion, mitochondrial dysfunction and damage, nuclear DNA damage and necrosis – so the overall physiology of the response is fairly similar between humans and mice. There is also sterile inflammation in people as well as in mice, but whether it amplifies liver injury or is in response to the damage and only responsible for clearing the cellular debris is a matter of controversy (Jaeschke et al., 2012b).

6.5 Making safer new drugs without animal testing – Paul Watkins

Among the kinds of drug-induced adverse events not predicted well by animals are cardiovascular and hepatic toxicities, which are the main causes of drug development failure at all stages and for regulatory actions on drugs post approval. Animals can be much more susceptible than humans to certain drug toxicities and therefore raise false safety concerns. For example, even a few tablets of ibuprofen can cause severe liver and kidney damage to a dog. Hence, ibuprofen would never have progressed to human trials had the dog been used in preclinical testing, yet ibuprofen is among the safest drugs in clinical use today. Conversely, there are many examples where drugs shown to be entirely safe in preclinical animal studies can end up causing severe adverse drug events in patients.

The main causes of hepatotoxicity in patients are “off-target” (i.e., not related to the drug’s intended action) and “idiosyncratic” (i.e., resulting from individual susceptibilities). With idiosyncratic toxicity, the vast majority of people can take a drug and it will be totally safe, but a small fraction of individuals will respond badly.

In terms of minimizing animal sacrifice in drug development, the first question is: can we come up with *in vitro* cell systems that would be as good as *in vivo* animal testing? Most scientists familiar with the problem would probably agree that we can. The DILISym™ (drug-induced liver injury simulation) initia-



tive is a public-private partnership with pharma industry partners contributing money, previously unreleased data and intellectual capital to generate a system for use in modeling DILI. DILIsym™ version 1a includes information for modeling several different species, including mouse, rat, dog and human. The modeling approach taken is “middle out,” starting with the toxic event at the organ level and modeling towards specific pathways and biological processes only as far as necessary to explain the species differences in toxicity susceptibility. The model can also take into account inter-individual variability. The goal is for DILIsym™ to be a routine tool to guide first-in-humans dosing and interpretation of DILI mechanistic biomarkers by January 2015. Because DILIsym™ utilizes data obtained largely from *in vitro* systems, we believe DILIsym™ will substantially limit the use of animals in preclinical safety testing.

The second question is can human *in vitro* systems and modeling represent an advance over current preclinical safety testing and lead to the complete elimination of animals in the safety testing of new drugs? An important effort in this regard is the Drug Induced Liver Injury Network (to continue funding by the National Institutes of Health until 2018) which finds and enrolls in a registry the rare patients who have experienced clinically important “idiosyncratic” DILI due to any drug. Biological samples, including DNA, are collected and archived from each subject. To identify genetic risk factors underlying susceptibility to DILI, we have performed a genome-wide association study (GWAS) on the entire cohort of DILI cases, hoping to find general susceptibility factors of all cause DILI. We did not find significant associations; however, it may still be the case that there are common risk factors associated with specific subgroups of drugs and a variety of studies are underway at the Hamner Institutes to “cluster” drugs for GWAS analysis by common effects on adverse outcome pathways in cultured human cells. An interesting finding is that when sufficient cases of DILI due to a single drug is subjected to GWAS, only HLA alleles have been identified as risk factors. The most impressive association to date is with the antibiotic flucloxacillin: if you have one specific HLA allele you have up to an 80-fold increased risk of developing liver toxicity due to this drug. These observations mean we have to add adaptive immune responses (between adverse outcome pathways and cell death) to the models describing at least some mechanisms underlying idiosyncratic liver toxicity.

Each of the subjects in the DILIN registry have agreed to be recontacted and offered enrollment in additional studies, and it has therefore been possible to develop liver-like cultures using induced pluripotent stem cells (iPSCs). The goal will be to prepare these cultures from DILIN subjects and compare responses to the implicated drug with responses in similar culture systems prepared from individuals tolerant to the drug. The objective will be to observe differences between cases and controls in the initiation of adverse outcome pathways that should inform early events that are necessary, but may not be sufficient, to cause DILI. Identification of these events should lead to improved *in vitro* systems to assess liver safety of new drugs in man.

In summary, cellular systems, including iPSC-derived, liver-like culture systems prepared from DILI patients, together with

in silico modeling approaches have the potential to reduce or eliminate animal testing while improving safety prediction during drug development.

6.6 Virtual Liver: Predicting Liver Injury – Imran Shah

From the information on 544 environmental chemicals in EPA’s Integrated Risk Information System (IRIS) database, human health risk assessments for 20% are based on liver toxicity. The liver is a complex, dynamic system and not simply a bag of enzymes. How do you assemble the pieces in a meaningful way so that we can predict toxicity? One objective is to come up with screening tools that could begin to point out and connect MIEs that lead to adverse downstream effects; however, there are many pathways to any given adverse effect. So, perhaps a more fruitful approach is to start with the histopathology and work backwards to the MIE.

A main objective of the high throughput toxicology approach being taken by the Tox21 projects is the identification of reproducible biomarkers; however, this is a very difficult problem in the absence of biological context. If we cannot do this in an entirely empirical way, maybe we can approach it through computational modeling. The Virtual Liver approach involves simulating different aspects of the toxicological process, modeling context by including exposure, ADME (using QSAR and *in vitro* assays), and cellular and histological effects to describe AOPs.

One question to ask in describing an AOP for a given set of effects (e.g., liver cancer) is how many chemicals are there that produce this set of effects in two different species, for example in both mice and rats (Shah et al., 2011). To prioritize chemicals for hepatocarcinogenic assessment, we can use the degree of association between nuclear receptor activity seen in humans and the extent of cancer progression you see in rodents.

In terms of modeling, we are currently “stuck” trying to use cellular effects to model hepatocyte necrosis, apoptosis and proliferation. The Virtual Liver approach begins with simulating internal exposure and from there molecular, cellular and *in vivo* organ-level effects. Here is a key distinction from all the other approaches you will hear about: we do not assume there is a particular kind of lesion *a priori*. Injury is not defined by the model. We look at the initial conditions and we run the simulation to model how a normal lobule would react to chemical exposure. To get the variability inherent in human beings, we generate different models. The Virtual Liver model allows the evaluation of alternative hypotheses.

To model a human 28-day repeated dose exposure, the Virtual Liver Framework Consortium takes the approach of simulating the concentration of the chemical in liver cells using *in vitro* clearance and partitioning, simulating molecular, cellular and tissue responses to chemicals using ToxCast™ *in vitro* assays, and estimating the hepatic response at 28-days using *in silico* modeling. The specific parameters were derived using *in vivo* data adjusted by applying molecular logic informed by pathway knowledge. The success of this approach depends on a systematic approach to compiling information from multiple sources; for example, approximately 200 papers describing phenobarbi-



tal toxicity were used to link molecular events into a plausible pathway hypothesis.

In the question-and-answer following this paper, it was noted that AOPs are dose- and time-dependent; accounting for context is critical and connecting the dots must be done in a careful way, for example, based on Bradford Hill criteria. It was also pointed out that dose-response is important in discriminating between different possible outcomes for the same chemical. It is important to dig more deeply into the actual perturbation that leads to the toxicity. The omics technologies are capable of providing enormous amounts of data; however, the idea that more information must surely be better is tempered by consideration of signal-to-noise; knowing what information is relevant is critical.⁹

6.7 An integrated predictive platform for DILI: Combining *in silico* and *in vitro* methods to predict toxicity mechanisms *in vivo* – Kas Subramanian

The approach of Strand Life Science's Virtual Liver project is to build a stable, normal liver model *in silico* (Strand has developed a rat version and is working on a human version). *In vitro* assays measure the effect of a drug/metabolite at certain points in the system; they do not measure toxic chemical endpoints. Prediction of toxicity is a combination of assays and computational models.

The idea is to reveal critical pathways and mechanisms of toxicity. The utility of the model is to aid drug design; to maintain potency while minimizing toxicity. The model is not a database of interactions, and it is not a QSAR model. The virtual liver consists of processes, pathways, etc., that are quantitatively linked and allows one to run simulations, make predictions and develop insights into the toxicity process. It is a model of biology; it is not "trained" using existing drugs or chemicals, and therefore is not limited by a chemical space (but rather is limited by our knowledge of biological space). The virtual liver can model hepatotoxicity, metabolic disease and cardiovascular disease.

The Virtual Liver project takes a top-down, integrated approach, similar to completing a jigsaw puzzle by piecing together the outline first: the boundary is the phenotype. Experimental human gene expression and other kinds of data start filling in the center part of the puzzle, and the gaps are filled by hypothesis. The process involves identifying key processes and writing a series of differential equations to model each state of the network. This approach does not look for initiating events.

To validate the model, the first question is to ask whether it represents normal physiology. If it does, then additional simulations test whether the model correctly predicts the effects of environmental insults (drugs or toxicants) or genetic mutations (the model is for acute injury and does not account for changes over time that are involved in chronic insults). Over 40 compounds have been tested in the model with reasonable predictive accuracy and mechanistic insight.¹⁰ An important aspect of the model is the ability to look at combinatorial effects of differ-

ent networks; if you do not take into account what effects other pathways have on a particular pathway you might be missing something. This model also has the capacity to superimpose the effects of diet, diabetes, etc. and can provide insights into idiosyncratic events.

The project was started in 2005. Over the last two years many of the assays have been identified and optimized and are ready to test a batch of chemicals. The project has had support from a grant from the World Bank and the Indian government's department of biotechnology.

In the discussion following this presentation, a questioner asked whether the model is changed when the predictions are not correct. There could be two reasons for incorrect predictions: the available assays may not be capturing a key event (the Strand team estimates only about 60% of the necessary assays are in place), or the model is predicting something other than is being observed; in that case, the modelers have to ask if they have the biology right or need to go back and add more information. The idea was raised that the model is essentially asking "at what point will the airplane crash?" You want to be able to ask your model what kinds of things go wrong and where, and verify that with experimental data. A step-by-step analysis has begun for some chemicals (e.g., troglitazone), but a systematic approach to this question has not yet been performed.

6.8 *In vitro* approaches for predicting liver specific toxicity – Colleen Toole

Predicting liver toxicity *in vitro* requires understanding hepatic physiology and function as well as things like route of exposure and blood flow. It also requires the proper tools, including cell models, that represent key processes and biomarkers that monitor liver functions; pathways are critical to tying these together. Orally administered chemicals may be metabolized by the gut prior to arriving in the liver; chemicals that are administered by other routes may also be modified before entering the liver. In addition, CYP metabolism plays a large role in liver toxicity – any *in vitro* assessment methods need to take all of these things into account.

Chemical disruption of normal liver function can – but does not necessarily – lead to adverse effects. Chemicals that disrupt glycolysis or oxidative metabolism are among those that can cause liver toxicity. For example, the mitochondria in hepatocytes are the most sensitive targets for rotenone.

Modification by CYP enzymes can either activate or deactivate a chemical's capacity for hepatotoxicity. For example, in the case of diclofenac, a non-steroidal anti-inflammatory, there were a couple hundred cases of liver failure resulting in death. It only occurred when diclofenac reached a certain level and is thought to be caused by a reactive intermediate metabolized via CYP3A4 metabolism. Another intermediate produced via CYP2C9 metabolism is not as reactive. A drug (co-administered) that would inhibit CYP3A4 would reduce the toxicity of diclofenac.

⁹ Dr Shah recommended reading *The Signal and the Noise: Why So Many Predictions Fail – But Some Don't* by Nate Silver (Penguin Press, 2012).

¹⁰ An example of the response to cyclosporine A was presented.



Most drugs that inhibit the bile salt export pump (BSEP; in the canicular membrane) with an IC₅₀ value $\leq 25 \mu\text{M}$ are associated with liver toxicity (Kis et al., 2012). Inhibition of bile flow can lead to chronic liver toxicity.

Selection of an appropriate cell model is important when organ specific toxicity is the goal: primary hepatocytes from the species of interest are preferred – e.g., dogs, humans. Questions to ask of the model include whether it is capable of intrinsic clearance, metabolism, transport and/or polarization (Wilkening and Bader, 2003). Primary hepatocytes grown in sandwich cultures (e.g., a collagen Matrigel sandwich) that can form canicular membranes begin to look like true hepatocytes and address these requirements fairly well. Monitoring multiple endpoints of general cell health as well as liver-specific functions provides the most accurate prediction of liver toxicity. When exposure is by a route other than oral, an additional cell type that can mimic the route-dependent metabolism is needed; CeeTox sets up multiple cell models for determining organ specific toxicity.

In summary, *in vitro* models can identify liver specific toxicity; cell models should be species-specific and should possess key functions of the liver; *in vitro* data must consider pharmacokinetic data; and *in vitro* endpoints should measure both liver specific function and general cell health.

Among the points raised in discussion of this presentation were: (1) assessing the metabolic competence of all of these cell systems is critical, and (2) how do you deal with metabolites created in one tissue that circulate to another? Generally, metabolites that are in the liver are excreted through the bile; where it becomes an issue is when the liver is failing so that physiologic steps of pharmacologic clearance are not happening properly. The metabolites of some chemicals (e.g., aromatic amines) are metabolized by the liver and can affect other organs (e.g., kidney).

6.9 Description of prototype Modes-of-Action related to repeated dose toxicity: Overview of fibrosis and steatosis pathways – Brigitte Landesmann

In the context of SEURAT-1, prototype MoAs related to repeated dose liver toxicity have been developed (Landesmann et al., 2012). The working definition of MoA used for this work was taken from Boobis et al. (2008): “the sequence of key events starting with the interaction of an agent with a target cell, through functional and anatomical changes resulting in cancer or other adverse health effects.”

A stepwise approach according to OECD guidance for AOP development has been applied: the first steps were the selection of the AO and the MIE, the two anchors that also specify the AOP, followed by the study of the relevant physiology that is underpinning the process. This step is crucial because AOP development requires knowledge of the underlying biological mechanisms to understand how sufficient perturbation of normal function by a chemical may result in dysregulation and, ultimately, toxicity. Next, the intermediate events were identified and, subsequently, the elaborated pathway was graphically presented. Then the scientific evidence underlying the AOP had to be evaluated and reported by assessing the weight-of-evidence (WoE) according to the Bradford Hill criteria.

The selection of liver fibrosis and liver steatosis as AOs was due to the fact that they typically result from chronic injury and accord with SEURAT-1 research activities. Protein alkylation was chosen as the MIE for fibrosis because it is common to the two SEURAT reference chemicals for liver fibrosis (allyl alcohol and carbon tetrachloride). Nuclear receptor (NR) binding was chosen as the MIE for steatosis, because endocrine disrupting chemicals (EDC) also act through interaction with NRs and there is globally increased attention in this area in relation to EDC. Liver X receptor (LXR) was chosen from the six relevant NRs because the LXR agonist T0901317 is one of SEURAT’s reference chemicals for liver steatosis. A systematic literature search was undertaken with emphasis on key studies and review papers and consecutive analysis according to different levels of biological organization. This process is time-consuming and expertise is needed to choose the relevant level of detail, as well as to resolve differences between studies (e.g., species, cell model, concentrations). A diagram delineating the pathway was created, and the pathway was evaluated using Bradford Hill criteria for a WoE approach, including an evaluation of the number and quality of the available studies, the coherence and consistency between the findings, the biological plausibility, and the relevance of the used models and consideration of alternative explanations. Finally, the pathway development was reported according to OECD guidance.

Some problematic issues important to consider when developing MoAs are: the level of detail in data collection; linearity vs. feedback and feed-forward loops and complex inter-relations; how to identify key events; how to incorporate quantitative evaluation (dose-response relationships) and temporal relationships.

In the liver fibrosis MoA being described here, the MIE is alkylation caused by allyl alcohol and carbon tetrachloride and the supporting evidence was judged to be “well established” (noting that this is not the only possible MIE – any insult that results in cell death can initiate fibrosis). The intermediate events identified were hepatocyte injury causing necrosis or apoptosis (with “very strong” scientific evidence), activation of hepatic macrophages (Kupffer cells) (very strong), TGF- β 1 expression (very strong), stellate cell activation (very strong), collagen accumulation (very strong), oxidative stress (strong), and chronic inflammation (very strong). Many organs in addition to liver are subject to fibrosis (e.g., lung, kidney, heart, skin, pancreas), and there is some evidence for common preserved pathways across species.

Likewise, steatosis was described as in the introduction, with the MIE being LXR activation. Intermediate events at the molecular level include: auto-regulation of the LXR α (up regulation, positive feedback); increase in expression and activity of the carbohydrate response element binding protein (ChREBP); increase in expression of the sterol response element binding protein 1c (SREBP-1c) from LXR activation and from the ChREBP; induction of lipogenic enzymes from the SREBP-1c, up-regulation of the free fatty acid uptake transporter FAT/CD36, induction of the fatty acid synthase (FAS) and induction of the stearoyl-CoA desaturase 1 (SCD1). These intermediate events lead to *de novo* synthesis of fatty acids and triglycerides and fat influx from the peripheral tissues to liver.



Intermediate events at the organelle level include mitochondrial toxicity attributed to increased lipid peroxidation and reactive oxygen species overproduction and endoplasmic reticulum stress as a consequence of increased hepatic lipids. Intermediate events at the cellular level include cell burst (fat accumulation), cell cycle arrest (followed by failure of compensatory cell division) and necrosis and/or apoptosis (resulting from fat accumulation causing oxidative stress, energy depletion, cytokine release, and/or mitochondrial toxicity). Intermediate events at the tissue level include fat accumulation (either microvesicular or macrovesicular) and finally liver steatosis (when liver fat reaches greater than 5-10% of the liver by weight).

The elaborated pathway was evaluated according to OECD guidance: 1) concordance of dose-response relationships: no available dose-response curves but sporadic numerical data; 2) temporal concordance among the key events and adverse outcome: the sequence of the events is in strong agreement; 3) strength, consistency and specificity of association of adverse outcome and initiating event: strong and consistent association but not specific; 4) relevance: LXR expression and all the other elements are considered adequately conserved from rodents to humans; 5) uncertainties, inconsistencies and data gaps: there are differences in LXR expression between cell lines; the MoA could be affected by ethnicity and the life stage of the individual; there are possible effects by factors related to the formation of steatosis including trends in adipose tissue deposition, total body fat, and visceral and subcutaneous adipose tissue (which vary among different life stages, sexes and ethnicity). The biological plausibility, coherence and consistency of the experimental evidence and the existence of alternative mechanisms have not yet been described.

7 Discussion: Pathway approaches to liver toxicity

Discussion for this session was led by Imran Shah, following a set of questions prepared by the organizing committee. The questions included: How generalizable is the pathway in terms of chemical space; how central is the particular described MIE to the observed adverse outcome; what uses can be made of the pathway at its current level of completeness; and what elements are missing that are necessary for other applications.

7.1 The fibrosis pathway

While fibrosis is a general response that can be caused by many and varied MIE, this particular MIE may be fairly specific to the chosen chemicals (allyl alcohol and carbon tetrachloride) (see Figure 4, Willett et al., 2014). There are several other MIEs leading to fibrosis, for example, alcohol (ethyl alcohol) is a major cause of fibrosis but is not a protein binder and is not included in this MIE. In general, initial key events are chemical-specific, whereas later intermediate events are disease-specific. If apopto-

sis is a key event in fibrosis, then any MIEs that lead to apoptosis should also lead to fibrosis, but in fact, not all chemicals that cause apoptosis lead to fibrosis. For this particular MIE, the rate of apoptosis is the important factor; therefore, what is really needed is the dose and temporal data. To inform whether a chemical will cause liver fibrosis in contrast to other apoptosis-related adverse outcomes, it must be clear that the chemical of interest actually reaches the liver (ADME information could be integrated as part of the AOP). Further characterization of intermediate events is needed and is necessary for identification of key events.

It was agreed that this MIE is fairly well described and can be used for read-across, and that liver fibrosis is a good anchor for an endpoint because it relates to something that can be looked at in a relevant way (that is, the endpoint can be observed in both an animal model and clinically). Participants agreed that aiming for a perfectly-described AOP is neither possible nor necessary. What is needed is sufficient information to make a particular decision. AOPs are plausible abstractions that are refined in increments, but even less-sophisticated or less-refined AOPs can be useful in the decision-making process.

The current fibrosis AOP as described is a useful starting point for capturing relevant intermediate events. It allows investigators to do targeted testing, determine missing information, and then develop a series of hypotheses that can be tested within a limited amount of time. Also important for the utility of this AOP is to resist thinking of fibrosis as a binary event. Histologically, normal livers contain many of the features of fibrosis; the diagnosis is a matter of degree.¹¹

The most important aspects of proving causal relationships according to the Bradford Hill criteria are temporal and dose-response relationships. Temporality has been shown for the main intermediate events in this (fibrosis) MoA, but the dose-response data are lacking (Landesmann et al., 2012). In addition, it will be important to compare to chemicals which cause alkylation but not fibrosis, as this could help distinguish the specificity of the MIE and could help in making quantitative distinctions in thresholds that result in adaptation versus adversity.

7.2 The steatosis pathway

For the adverse outcome of steatosis, there are potentially several competing pathways, including LXR and multiple nuclear receptor-dependent pathways (see Figure 5, Willett et al., 2014). This MIE is likely only one of several that can lead to steatosis. For example, toxicogenomic studies have been done to identify genes that are specifically activated by steatosis, including studies testing the effects of agents in microarrays followed by validation by PCR. In these studies, CD36 was not one of the genes identified, so CD36 up-regulation may be specific for this chemical and may indicate that this MIE is not generalizable to other chemicals; however, the link between CD36 and steatosis was only recently recognized, so future studies may discover a broader applicability

¹¹ Liver fibrosis is difficult to diagnose clinically, especially at early stages. Only recently new serum markers and non-invasive techniques such as ultrasound-based and magnetic resonance elastography have been developed. The backbone for diagnosis is still the biopsy. Fibrosis stages are classified semi-quantitatively by both the quantity and location of the fibrosis. Prevalence of advanced fibrosis was found to be around 2.8% in a French general population aged 40 years or older. Several independent risk factors have been identified – but early fibrosis is still hard to detect via non-invasive biomarkers (Poynard et al., 2010).



In addition, there are two different forms of steatosis, micro- and macrovesicular. Microvesicular steatosis is very serious, related to beta oxidation and mitochondrial toxicity, and does not necessarily progress to macrovesicular. Macrovesicular steatosis is associated with many physiological syndromes, including alcohol consumption, and involves suppression of fatty acid/triglyceride efflux. There are some differences in the composition of the vesicles in micro- and macrovesicular steatosis (fatty acid vs triglyceride). Perhaps this AOP should be split into micro/macro steatosis for the tissue endpoint. A steatosis expert might be needed to make some of these distinctions.

The described pathway is a good start for the biological description of steatosis, but may be too narrowly applicable and insufficiently developed to be used in its current form.

8 Using pathway based approaches to inform decision-making

This discussion was facilitated by Mel Andersen, and addressed issues such as relating *in vitro* observations to *in vivo* adversity, how to identify assays that can model intermediate or key events, and how to make sure all mechanisms of liver toxicity are sufficiently covered, including MIEs and AOs for different chemical categories.

To use pathways as a construct to inform decisions, an initial step is to describe the causal linkages between the intermediate events (a “qualitative AOP”), and then begin to quantify these relationships. This is an iterative process. As noted earlier, different uses of AOPs require different levels of certainty and quantification. For risk assessment, quantification is necessary: how much do you need to perturb a given step in the pathway to provoke the next step? To differentiate adverse conditions from perturbations within homeostatic bounds, thresholds will need to be defined to separate, for example, endocrine disruptors from non-endocrine disruptors. Once the relationships have been quantified, it may be possible to identify an “upstream event” as a quantitative point of departure for a downstream AO.

Discussants pointed out that adversity is a matter of degree: for example, clinically elevated enzymes are not considered alarming until bilirubin levels rise (at which point it is indeed serious). An AOP needs to reflect a grading system of severity that maps to existing clinical guidelines.

This process is resource-intensive, so we want to first focus resources on investigating those endpoints that are legislatively mandated and are actually driving the risk assessments.

It is likely that different assays are needed to build the pathways than are needed in integrated testing strategies using the pathways to evaluate chemicals. A related question is what assays to use to quantify the relationships. For example, ER binding and gene expression are two key events in the estrogen receptor-mediated reproductive impairment AOP: so now we need to do directed/targeted testing to make the quantitative links between them.

We are just beginning to get glimpses of the complex system that leads to different forms of liver toxicity, so imposing lin-

earity on the AOP in these early stages makes it more manageable. We must start small and accept some uncertainty, knowing that we will be adding information and complexity step-wise until all observations are accounted for. Understanding the full context of an AOP is a little bit like moving traffic: there may be one well-established route to our destination, but in certain events – an accident, for example – we want to know the alternate routes to that destination. In fact, a vision for the future would be to build a collection of fibrosis pathways, for example, with different MIEs. Then someone can go to the Toolbox, ask which pathways lead to fibrosis, and use that information in a diagnostic context to identify target tests for an individual exhibiting fibrosis. This vision underscores the importance of creating a centralized knowledge base. As one participant noted, a strength of the AOP approach is that you can take information from many different sources and ultimately bridge the needs of the different communities.

9 Concluding remarks

The participants in this workshop discussed the state-of-the-art in adverse outcome pathway development and focused discussion on two pathways in particular, the prototype modes-of-action for liver fibrosis and steatosis developed by the European Commission Joint Research Centre. There was consensus that the AOP concept is extremely useful for organizing information at all levels of biological organization (from the molecular, cellular, tissue, to the population-level) and across the spectrum of biological information (chemistry, bioinformatics, omics, histology, guideline animal data). The OECD definition of the AOP as a linear portion of a biological network from a single molecular initiating event (MIE) to a single adverse outcome (AO) is useful theoretically as a simplifying concept, but may be too limiting to be practical; some AOPs may necessarily involve branching. For many participants, the terms Mode-of-Action (MoA) and AOP were synonymous; for some, the AOP encompasses the MoA, which is the upstream portion of the AOP, from the MIE through to resulting cellular and tissue changes. There was consensus that the quality assurance methodologies described in OECD guidance (e.g., to assess data quality, causal relationships and completeness of the AOP) were appropriate and essential. The confidence and completeness of an AOP will determine its usefulness: less complete AOPs can be useful for chemical categorization and prioritization of chemicals for further testing while AOPs with high confidence and quantitative understanding of the links between events are necessary if they are to be used to inform risk assessment. There was also consensus that the best way to progress is to develop case studies; indeed, several working groups are taking this approach. In addition, focusing the more theoretical conversation about AOP development on a specific AOP is extremely helpful; for example, many general concepts were discussed focusing on the fibrosis AOP.

Liver toxicity is an appropriate subject for AOP case studies as it is central to acute, subchronic and chronic chemical exposure, whether it be during repeated exposure to environmental



chemicals or to drugs, where liver toxicity is a leading cause of removal from the market. Drug-induced liver injury (DILI) is rare, making it hard to detect in clinical trials, but often fatal. DILI can also be highly dependent on the patient's own genetic makeup and physiological condition – underscoring the need for a better understanding of the underlying biology to better predict the occurrence and outcomes. While there are described diagnostic markers for DILI, these are not specific and, more importantly, are not predictive of future occurrence. Understanding liver toxicity requires an in-depth understanding of normal liver function as well as the pathway perturbations that occur during toxicity. Pathway-based approaches to liver toxicity and DILI are important for identifying predictive biomarkers.

The liver fibrosis MoA presented by the JRC was delineated using information from exposure to allyl alcohol and carbon tetrachloride, and therefore the MIE is protein alkylation. While this MIE is specific for these two chemicals and is not applicable to other toxicants (e.g., ethanol), experience with and information about these other well-studied chemicals can inform general understanding of liver biology and fibrosis MoAs. Conversely, it is important to begin with case studies and build pathways out from these well-described prototypes.

Steatosis (fatty liver) may be caused by several MIEs, one of which is interference with the array of nuclear receptors involved in homeostasis of fatty acid metabolism; the MoA discussed involved activation of the Liver X receptor (LXR) as the MIE. Binding to LXR causes an induction of transcription factors leading to de novo fatty acid and triglyceride synthesis as well as an influx of fat from peripheral tissues to the liver. Fatty acid accumulation can be protective and its appearance in humans is not necessarily cause for alarm; however, its appearance in animal studies during drug development may stop further development. Micro- and macrovesicular steatosis are separate clinical entities: microvesicular steatosis is more serious and indicates a problem with mitochondrial beta-oxidation and breakdown of fatty acids which affects the production of ATP. Macrovesicular steatosis is associated with many conditions, including alcohol-suppression of fatty acid/triglyceride

efflux. There are differences in the composition of the vesicles in micro and macro (fatty acid vs. triglyceride). It was suggested that this AOP should be split into two – one for micro- and another for macrovesicular steatosis.

Bioinformatics and modeling approaches are being used to identify predictive markers or signatures for liver toxicity; these data can also inform pathway development and the two approaches can be used to support predictive modeling in complementary ways: bioinformatics can help delineate MIE and upstream intermediate pathway events, and formal pathway descriptions can be used to support hypothesis-based testing that can be used to “validate” both the predictive markers and signatures, as well as the pathway assumptions. Bioinformatics and modeling approaches can also be used to incorporate the influences of genetic variability on normal and perturbed pathways. To fully inform pathway development, it is also essential to incorporate the knowledge from histologists and clinicians familiar with the biological outcomes at the tissue and individual levels. In the case of DILI, there are also large human databases that contain information that can be used to inform idiosyncratic factors.

Some important considerations for the future of AOP development in general and liver AOPs in particular that were discussed included: accounting for context – AOPs are species, tissue, timing and dose-dependent; consideration of metabolism, both before entering the liver and within the liver; the fact that more information is not always better – a consideration of signal-to-noise and knowing what information is relevant is critical; identification of appropriate tools to test integrated strategies including appropriate cell types and tissue constructs; incorporation of severity – prediction when the “tipping-point” of adversity has been reached (quantification of the relationships between intermediate effects is necessary); appropriate design of assays to inform the quantitative relationships between intermediate events; the importance of a centralized and accessible knowledge base. At the conclusion of the workshop, participants were asked to share what they saw as the most salient “take-away” message(s) from the presentations and discussions; these are compiled in Table 1.

Tab. 1: Final thoughts

Development
AOPs should be generic and applicable to chemically diverse compounds that induce a certain type of toxicity. Moreover, the relevance of AOPs should be substantiated by human-relevant clinical data as it is not a simple summary of mechanisms described in scientific literature.
AOPs express incomplete views of the sequential phenomena involved in linking exposure to outcome.
AOPs do not imply linear relationships between phenomena – the nonlinearity (feedback) is implicit.
Perspective on key events will differ according to who is assessing.
MIE: unlikely that one MIE leads to one IE and one AO, however, it provides a starting point for a more complex explanation.
<i>In vitro</i> and <i>in silico</i> approaches provide useful fingerprints of homeostasis and how it can be altered or disrupted by changing one or several parameters and which pathways could be targeted.
Need for language agreement, shared terminology.



Development (continued)
It is helpful to have a conversation between individuals from various backgrounds and learning their needs and perspectives, especially in a regulatory environment.
Excellent to discuss a prototype pathway (fibrosis) as it makes the discussion more real. Hopefully this can continue to the phase of “how do <i>in vitro</i> models fit into this AOP?”
Important to have medical professionals in the discussion on “adversity” – I really believe it is truly adverse events in humans that the AOP framework can help us with.
Use
AOPs are serving to facilitate communication of information on toxicology to regulators and governmental agencies, primarily in the EU. AOPs explicitly display our understanding of the relationship of specific processes (steps) to effects that serve as regulatory endpoints. More widespread use of these AOPs will be useful in designing targeted testing strategies, better short term test methods, and support initiatives that will eliminate/reduce reliance on animal testing for decision making.
The phrase “designed for purpose” is a perfect description of a use of AOPs – to help design targeted testing rather than test a whole set of assays likely to be unrelated to the hazard associated with your chemical.
“Form follows function” – The structure/design of the AOP is determined by its application.
The use of AOP should be clearly articulated, since experimental validation will depend on intended usage.
AOP does not necessarily reduce uncertainty in risk assessment, but does inform it.
End-use for the AOP needs to be defined as it relates to the type and nature (quantitative vs. qualitative) of information.
Needs for drugs and chemicals are different: for drugs we are looking for a needle in a haystack – for chemicals we need only to be sure for 95% of cases – and what may seem not certain enough for clinical trials in fact will be sufficient for estimating risk of chemicals.
An “established AOP” can be qualitative, based on a minimal set of data (causal linkages between key events established) – depends on the use.
How different people/groups will/can use AOPs in different ways is exciting and perhaps we could be more explicit about this.
Next steps
AOP concept is established – next step is to make use of the tools and create data.
Need activity on various fronts to encourage flexible and widely accepted frameworks for incorporation of AOP into regulatory context.
Examples of AOPs are key to progress the concept.
Populate AOPs with DATA – move from concept to application.
How do we implement the concept of “necessary but not sufficient” into AOP development and acceptance?
Identification of thresholds is needed – the importance of the <i>in vitro</i> dose (in the medium/intracellular) will be important in identifying a true threshold.
Multiple agencies with similar agendas must coordinate their efforts.
Quantification is critical – is where the rubber meets the road.
The shift from theoretical discussion to attempts to find a practical application yet still remain flexible to individuals’ needs and applications is critical to further the discussion.
A very stratified and transparent method for central data collection, assay development and AOP publication needs to be specifically identified and advertised.

References

- Andersen, M. E. and Krewski, D. (2010). The vision of toxicity testing in the 21st century: Moving from discussion to action. *Toxicol Sci* 117, 17-24. <http://dx.doi.org/10.1093/toxsci/kfq188>
- Andersen, M. E., Clewell, H. J., Carmichael, P. L. and Boekelheide, K. (2011). Can case study approaches speed implementation of the NRC report: “Toxicity testing in the 21st century: a vision and a strategy?” *ALTEX* 28, 175-182. <http://www.altex.ch/Current-issue.50.html?iid=127&aid=1> (accessed 20.02.2014).
- Andersen, M. E., McMullen, P. D. and Bhattacharya, S. (2013). Toxicogenomics for transcription factor-governed molecular pathways: moving on to roles beyond classification and prediction. *Arch Toxicol* 87, 7-11. <http://dx.doi.org/10.1007/s00204-012-0980-6>



- Bhattacharya, S., Zhang, Q., Charmichael, P. L. et al. (2011). Toxicity testing in the 21st century: defining new risk assessment approaches based on perturbation of intracellular toxicity pathways. *PLoS ONE* 6, e20887. <http://dx.doi.org/10.1371/journal.pone.0020887>
- Boobis, A. R., Cohen, S. M., Dellarco, V. et al. (2006). IPCS framework for analyzing the relevance of a cancer mode of action for humans. *Crit Rev Toxicol* 36, 781-792. <http://dx.doi.org/10.1080/10408440600977677>
- Boobis, A. R., Doe, J. E., Heinrich-Hirsch, B. et al. (2008). IPCS framework for analyzing the relevance of a noncancer mode of action for humans. *Crit Rev Toxicol* 38, 87-96. <http://dx.doi.org/10.1080/10408440701749421>
- Bottini, A. A. and Hartung, T. (2009). Food for thought ... on the economics of animal testing. *ALTEX* 26, 3-16. <http://www.altex.ch/Current-issue.50.html?iid=104&aid=1> (accessed 20.02.2014).
- Cooper, S., Khatib, F., Treuille, A. et al. (2010). Predicting protein structures with a multiplayer online game. *Nature* 466, 756-760
- European Commission (2005). REACH and the need for Intelligent Testing Strategies. Institute for Health and Consumer Protection, European Commission, Joint Research Centre, Ispra, Italy. <http://reach-support.com/download/Intelligent%20Testing.pdf> (accessed 02.01.2014).
- Jaeschke, H., McGill, M. R. and Ramachandran, A. (2012a). Oxidant stress, mitochondria, and cell death mechanisms in drug-induced liver injury: lessons learned from acetaminophen hepatotoxicity. *Drug Metab Rev* 44, 88-106. <http://dx.doi.org/10.3109/03602532.2011.602688>
- Jaeschke, H., Williams, C. D., Ramachandran, A. et al. (2012b). Acetaminophen hepatotoxicity and repair: the role of sterile inflammation and innate immunity. *Liver Int* 32, 8-20. <http://dx.doi.org/10.1111/j.1478-3231.2011.02501.x>
- Jaworska, J. and Hoffmann, S. (2010). Integrated Testing Strategy (ITS) – opportunities to better use existing data and guide future testing in toxicology. *ALTEX* 27, 231-242. <http://www.altex.ch/Current-issue.50.html?iid=121&aid=1> (accessed 20.02.2014).
- Hartung, T. (2009). Per aspirin ad astra... *ATLA* 37, 45-47. http://www.frame.org.uk/atla_article.php?art_id=1238&abstract=true (accessed 20.02.2014).
- Hartung, T., Luechtefeld, T., Maertens, A. et al. (2013a). Integrated testing strategies for safety assessments. *ALTEX* 30, 3-18. <http://dx.doi.org/10.14573/altex.2013.1.003>
- Hartung, T., Hoffmann, S. and Stephens, M. (2013b). Mechanistic validation. *ALTEX* 30, 119-130. <http://dx.doi.org/10.14573/altex.2013.2.119>
- Hoffmann, S., Hothorn, L. A., Edler, L. et al. (2012). Two new approaches to improve the analysis of BALB/c 3T3 cell transformation assay data. *Mutat Res* 744, 36-41. <http://dx.doi.org/10.1016/j.mrgentox.2011.12.002>
- Kis, E., Ioja, E., Rajnai, Z. et al. (2012). BSEP inhibition – in vitro screens to assess cholestatic potential of drugs. *Toxicol In Vitro* 26, 1294-1299. <http://dx.doi.org/10.1016/j.tiv.2011.11.002>
- Kleensang, A., Maertens, A., Rosenberg, M. et al. (2013). Pathways of Toxicity. *ALTEX* 31, 53-61. <http://dx.doi.org/10.14573/altex.1309261>
- Krewski, D., Acosta, D. Jr., Andersen, M. et al. (2010). Toxicity testing in the 21st century: a vision and a strategy. *J Toxicol Environ Health B Crit Rev* 13, 51-138. <http://dx.doi.org/10.1080/10937404.2010.483176>
- Kwekel, J. C., Burgoon, L. D., Burt, J. W. et al. (2005). A cross-species analysis of the rodent uterotrophic program: elucidation of conserved responses and targets of estrogen signaling. *Physiol Genomics* 23, 327-342. <http://dx.doi.org/10.1152/physiolgenomics.00175.2005>
- Landesmann, B., Goumenou, M., Munn, S. and Whelan, M. (2012). *Description of Prototype Modes-of-Action Related to Repeated Dose Toxicity*. JRC Scientific and Policy Reports 75689. <http://dx.doi.org/10.2788/71112>
- MacKay, C., Davies, M., Summerfield, V. and Maxwell, G. (2013). From pathways to people: applying the adverse outcome pathway (AOP) for skin sensitization to risk assessment. *ALTEX* 30, 473-786. <http://dx.doi.org/10.14573/altex.2013.4.473>
- McMullen, P. D., Bhattacharya, S., Woods, C. G. et al. (2014). A map of the PPAR α transcription regulatory network for primary human hepatocytes. *Chem Biol Interactions* 209, 14-24. <http://dx.doi.org/10.1016/j.cbi.2013.11.006>
- NRC – National Research Council (2007). *Toxicity Testing in the 21st Century: A Vision and A Strategy*. Washington, DC, USA: National Academies Press.
- OECD – Organization for Economic Co-operation and Development (2013). *Guidance Document on Developing and Assessing Adverse Outcome Pathways*. Paris, France: OECD. [http://search.oecd.org/officialdocuments/displaydocumentpdf/?cote=en/jm/mono\(2013\)6&doclanguage=en](http://search.oecd.org/officialdocuments/displaydocumentpdf/?cote=en/jm/mono(2013)6&doclanguage=en) (accessed 20.02.2014).
- Poynard, T., Lebray, P., Ingiliz, P. et al. (2010). Prevalence of liver fibrosis and risk factors in a general population using non-invasive biomarkers (FibroTest). *BMC Gastroenterol* 10, 40. <http://dx.doi.org/10.1186/1471-230X-10-40>
- Russmann, S., Kullak-Ublick, G. A. and Grattagliano, I. (2009). Current concepts of mechanisms in drug-induced hepatotoxicity. *Curr Med Chem* 16, 3041-3053. <http://dx.doi.org/10.2174/092986709788803097>
- Shah, I., Houck, K., Judson, R. S. et al. (2011). Using nuclear receptor activity to stratify hepatocarcinogens. *PLoS ONE* 6, e14584. <http://dx.doi.org/10.1371/journal.pone.0014584>
- Stephens, M. L., Barrow, C., Andersen, M. et al. (2012). Accelerating the development of 21st-century toxicology: Outcome of a Human Toxicology Project Consortium workshop. *Toxicol Sci* 125, 327-334. <http://dx.doi.org/10.1093/toxsci/kfr248>
- Stephens, M. L., Andersen, M., Becker, R. A. et al. (2013). Evidence-based toxicology for the 21st century: opportunities and challenges. *ALTEX* 30, 74-103. <http://dx.doi.org/10.14573/altex.2013.1.074>
- Sun, B., Ross, S. M., Trask, O. J. et al. (2013). Assessing dose-dependent differences in DNA-damage, p53 response and genotoxicity for quercetin and curcumin. *Toxicol In Vitro* 27, 1877-1887. <http://dx.doi.org/10.1016/j.tiv.2013.05.015>

Thomas, R. S., Philbert, M. A., Auerbach, S. S. et al. (2013). Incorporating new technologies into toxicity testing and risk assessment: Moving from 21st century vision to a data-driven framework. *Toxicol Sci* 136, 4-18. <http://dx.doi.org/10.1093/toxsci/kft178>

US EPA (2009). EPA Issue Paper: Draft Effects-based Expert System to Predict Estrogen Receptor Binding Affinity. Federal Register docket number EPA-HQ-OPP-2009-0322 (available at regulations.gov, EPA-HQ-OPP-2009-0322-0002).

Vinken, M., Landesmann, B., Goumenou, M. et al. (2013). Development of an adverse outcome pathway from drug-mediated bile salt export pump inhibition to cholestatic liver injury. *Toxicol Sci* 136, 97-106. <http://dx.doi.org/10.1093/toxsci/kft177>

Wilkening, S. and Bader, A. (2003). Influence of culture time on the expression of drug-metabolizing enzymes in primary human hepatocytes and hepatoma cell line HepG2. *J Biochem Mol Toxicol* 17, 207-213. <http://dx.doi.org/10.1002/jbt.10085>

Willett, C., Caverly Rae, J., Goyak, K. O. et al. (2014). Pathway-based toxicity: history, current approaches and liver fibrosis and steatosis as prototypes. *ALTEX* 31, 407-421. <http://dx.doi.org/10.14573/altex.1401283>.

Zhang, Q. and Andersen, M. E. (2007). Dose response relationship in anti-stress gene regulatory networks. *PLoS Comput*

Biol 3, e24. <http://dx.doi.org/10.1371/journal.pcbi.0030024>
Zhang, Q., Bhattacharya, S., Andersen, M. E. and Conolly, R. B. (2010). Computational systems biology and dose-response modeling in relation to new directions in toxicity testing. *J Toxicol Environ Health B Crit Rev* 13, 253-276. <http://dx.doi.org/10.1080/10937404.2010.483943>

Acknowledgements

The authors would like to thank Marilyn Principe (John Hopkins Center for Alternatives to Animal Testing) for logistical support.

Correspondence to

Catherine Willett, PhD
Director, Regulatory Toxicology, Risk Assessment and Alternatives
The Humane Society of the United States
2100 L Street NW Washington, DC 20037
Phone: +1 240 599 6785
e-mail: kwillett@humanesociety.org

8. Tierversuchstagung des Schweizer Tierschutz STS

ERSATZMETHODEN FÖRDERN TIERVERSUCHE ERSETZEN

Donnerstag, 18. Dezember 2014
Kongresszentrum Hotel Arte
Riggenbachstrasse 10, CH-4600 Olten
10.00 – 16.30 Uhr

Anmeldung:
Schweizer Tierschutz STS
Telefon: 061 365 99 99
www.tierschutz.ch
Tagungsgebühr:
(inkl. Verpflegung und Tagungsunterlagen)
Vollzahler(in): 180.- CHF
Student(in): 90.- CHF

Tagungssprache: Hochdeutsch
Simultandolmetschung: Deutsch-Französisch und Französisch-Deutsch
Als Weiterbildungstagung [1/2 Tag] für das Fachpersonal für Tierversuche von der Vereinigung der Schweizer KantonstierärztInnen (VSKT) anerkannt.



SCHWEIZER TIERSCHUTZ STS

Ersatzmethoden fördern – Tierversuche ersetzen:
Zeit für einen Umbruch

Dr. med. vet. MLaw Julika Fitzl-Rathgen

Experimentation animale: Le mauvais modèle

Dr. chem. Isabelle Chevalley

Tierversuche und Ersatzmethoden:
Die Anforderungen des 3R-Postulates (Replace, Reduce, Refine)

Dr. med. vet. PHD Heinrich Binder

Validierung und Implementierung von Ersatzmethoden bei Industriechemikalien – Aktivitäten der Schweiz in einem internationalen Umfeld (OECD Testguidelines, REACH)

Dr. Markus Hofmann
Dr. Chantra Eskes, Eng, PhD

Praktische Anwendung von Alternativmethoden zum Tierversuch: Ein ganz persönlicher Rückblick über 30 Jahre

Dr. Toni Lindl

Animal Welfare und 3R – Perspektiven der Pharmaindustrie

PD Dr. Birgit Ledermann

Möglichkeiten und Grenzen der 3D-Gewebemodelle als alternative Testmethoden

Prof. Dr. Ursula Graf-Hausner

Ersatzmethoden an Universitäten

Prof. Dr. Pierre Cosson