Supplementary material: extended online version

Workshop Report: Developing Microphysiological Systems for Use as Regulatory Tools

Melvin E. Andersen[1], Kellyn Betts[2], Yvonne Dragan[3], Suzanne Fitzpatrick[4], Jesse L. Goodman[5], Thomas Hartung[6], Jonathan Himmelfarb[7], Donald E. Ingber[8], Abigail Jacobs[9], Robert Kavlock[10], Kyle Kolaja[11], James L. Stevens[12], Dan Tagle[13], D. Lansing Taylor[14] and Douglas Throckmorton[9]

[1]Hamner Institutes for Health Sciences, Research Triangle Park, NC, USA; [2]Freelance Science and Technology Writer, Takoma Park, MD, USA; [3]AstraZeneca, Waltham, MA, USA; [4]Center for Food Safety and Applied Nutrition, Food and Drug Administration, Silver Spring, MD, USA; [5]Food and Drug Administration, Silver Spring, MD, USA; [6]Johns Hopkins University Center for Alternatives to Animal Testing, Baltimore, MD, USA; [7]Kidney Research Institute, University of Washington, Seattle, WA, USA; [8]Wyss Institute for Biologically Inspired Engineering, Harvard University, Cambridge, MA, USA; [9]Center for Drug Evaluation and Research, Food and Drug Administration, Silver Spring, MD, USA; [10]U.S. Environmental Protection Agency, Research Triangle Park, NC, USA; [11]Cellular Dynamics International, Inc., Madison, WI, USA; [12]Lilly Research Laboratories, Greenfield, IN, USA; [13]National Center for Advancing Translational Sciences, National Institutes of Health, Bethesda, MD, USA; [14]Drug Discovery Institute, University of Pittsburgh, Pittsburgh, PA, USA

Summary

In the last few years, scientists have made progress in developing systems using human cells to test the effects of drugs and other substances. These systems are known as microsystems, microphysiological systems, or organs-on-a-chip. They have the potential to improve toxicity testing beyond currently available tools and to reduce the number of animals used. On May 10, 2013 scientists in academia, industry and regulatory agencies met in person and online to discuss the essential elements needed to develop these systems for use as regulatory tools, as well as pathways to their qualification. The one-day workshop was co-sponsored by the Food and Drug Administration, National Institutes of Health, National Institute for Environmental Health Sciences, National Center for Advancing Translational Science, Environmental Protection Agency, Johns Hopkins School of Public Health's Center for Alternatives to Animal Testing and the International Consortium for Innovation and Quality in Pharmaceutical Development.

Keywords: microphysiological systems, organs-on-a-chip

1 Introduction

In opening the workshop, Jesse L. Goodman, M.D., M.P.H., the chief scientist of the Food and Drug Administration (FDA) characterized microphysiological systems as combining computer science, systems biology, cell biology and modeling "to revolutionize tools and develop medical products… We all are here because we can see the tremendous possibility in improving what we in FDA think of as 'predictive science'" to help scientists predict the safety and effectiveness of products. He said that he felt that the science of toxicology has been underestimated and underappreciated. The new tools have the potential to be used in place of conventional animal testing.

Goodman predicted that microphysiological systems may impact both the specificity of toxicology and how it is modeled. "We have many compounds and interventions that we have to reject now that if we can improve specificity we may not have to reject." The tools may help illuminate the reasons for toxicity, as well as how genetics may come into play, he says. The tools also hold promise for furthering the study of disease models, Goodman said. "Being able to have 3-dimensional complex models using human cells offers tremendous potential."

Goodman noted that he often has to remind scientists that "a human is not a mouse." He pointed out that many drugs and treatments that look promising in murine models end up not being effective in humans for a variety of reasons that are often

http://dx.doi.org/10.14573/altex.1405151s

surprising. At the same time, he stressed: “a chip is not a human.” He also predicted that microphysiological systems may have to merge with stem cells, modeling, and systems biology into what he called a “pool or source of information” in order to be practical. The challenge for developing these new tools for use in regulatory science is to construct the science that shows where the tools are predictive – and where they are not.

2 Academic, industry and regulatory perspectives on the new tools

The first session of the conference featured talks by experts in academia, industry and regulation. Each gave their perspective on whether microphysiological systems had real promise or were little more that ‘pie in the sky.’

2.1 Academic perspective

First up to the podium to provide the academic perspective was Donald E. Ingber, M.D, Ph.D., the director of the Wyss Institute for Biologically Inspired Engineering at Harvard University. He is also the Judah Folkman Professor of Vascular Biology at Harvard Medical School and the Vascular Biology Program at Boston Children’s Hospital; and professor of bioengineering at the Harvard School of Engineering and Applied Sciences.

Ingber’s 4-year-old institute is involved in what he calls “high risk technology development,” and its faculty is organized by the problems it is attempting to solve. Wyss’ goal is to have products in the pipeline within 5 years, and it already has produced one product and has two clinical trials about to begin.

The institute’s projects include efforts to develop what it calls “biomimetic microsystems,” based on application of microfabrication approaches used to create integrated circuits, or chips, that contain living human cells that reconstitute organ-level functions. These microdevices are also known as “organs-on-chips.” They recreate tissue-tissue interfaces and physiological microenvironments critical to organ function and have the potential to expand the capabilities of cell culture models and provide low-cost alternatives to animal studies for drug screening and toxicology applications.

In 2010, Dongeun Huh, Ingber and others from Wyss and Harvard published an article describing what they call a lung-on-a-chip in *Science* (Huh et al., 2010). The chip recreates the alveolar-capillary interface, one of the major functional units in the lung and the site where oxygen enters the body. This same interface is where aerosol-based drugs are delivered, where some cancers can metastasize (Huh et al., 2010), and is a major site where pneumonias develop, among other things.

The human lung air sac is a relatively simple 3-dimensional structure, Ingber said. It consists of air overlaying a single epithelial cell layer that sits on a porous flexible extracellular matrix, which is then lined by another layer of capillary endothelium on its opposite side where it comes in contact with flowing blood. “A critical feature of the lung from a physiological perspective that is commonly ignored is that it is mechanically active,” he said. “Every time we breathe in and out there’s cyclic strain that deforms the entire tissue-tissue interface.”

The lung-on-a-chip is the size of a computer memory stick (Fig. 1), Ingber said. It was inspired in part by recent advances in microfluidics, which involves use of computer microfabrication techniques to construct networks of hollow channels that can control and manipulate fluids at very small, sub-millimeter, scales to take advantage of changes in how the fluids behave at these microscopic dimensions. The lung-on-a-chip has two adjacent hollow rectangular channels separated by a very thin 10 micron membrane containing 10 micron pores. The membrane is coated with extracellular matrix, and human lung sac epithelial cells are cultured on the top of the membrane, while human capillary endothelial cells are cultured on the opposite side of the same membrane to recreate the tissue-tissue interface of the lung air sac. Air is flowed over the epithelium and culture medium with or without human white blood cells is flowed through the vascular channel. To mimic physiological breathing motions, two larger chambers are fabricated into the device on either side of the channels with the tissue-tissue interface, and intermittent suction is applied. Because the device is fabricated from a flexible clear silicon rubber, the porous membrane and attached tissue stretch when suction is applied, and then recoil back when it is released, just like the cells of our air sacs do when we breathe in and out.

The researchers were able to use the lung-on-a-chip to mimic common lung functions and generate predictions about previously unknown functions that were confirmed in studies with whole mouse lungs. For example, they showed that exposure to airborne particulates in the form of colloidal silica nanoparticles can create an inflammatory response in the lung-on-a-chip, and they discovered that the cyclic mechanical strain of breathing accentuates the toxic and inflammatory responses. “Breathing alone does not achieve this, and the nanoparticles alone do not do it. Only together does this happen,” Ingber explained. The mechanical strain also increased the absorption of the nanoparticles from the airspace and into the blood fluid by 8 to 10-fold. “No one had ever seen this before,” he said, and when they tested this prediction in mice, the same thing happened.

The researchers were able to use the lung-on-a-chip to mimic common lung functions and generate predictions about previously unknown functions that were confirmed in studies with whole mouse lungs. For example, they showed that exposure to airborne particulates in the form of colloidal silica nanoparticles can create an inflammatory response in the lung-on-a-chip, and they discovered that the cyclic mechanical strain of breathing accentuates the toxic and inflammatory responses. “Breathing alone does not achieve this, and the nanoparticles alone do not do it. Only together does this happen,” Ingber explained. The mechanical strain also increased the absorption of the nanoparticles from the airspace and into the blood fluid by 8 to 10-fold. “No one had ever seen this before,” he said, and when they tested this prediction in mice, the same thing happened.

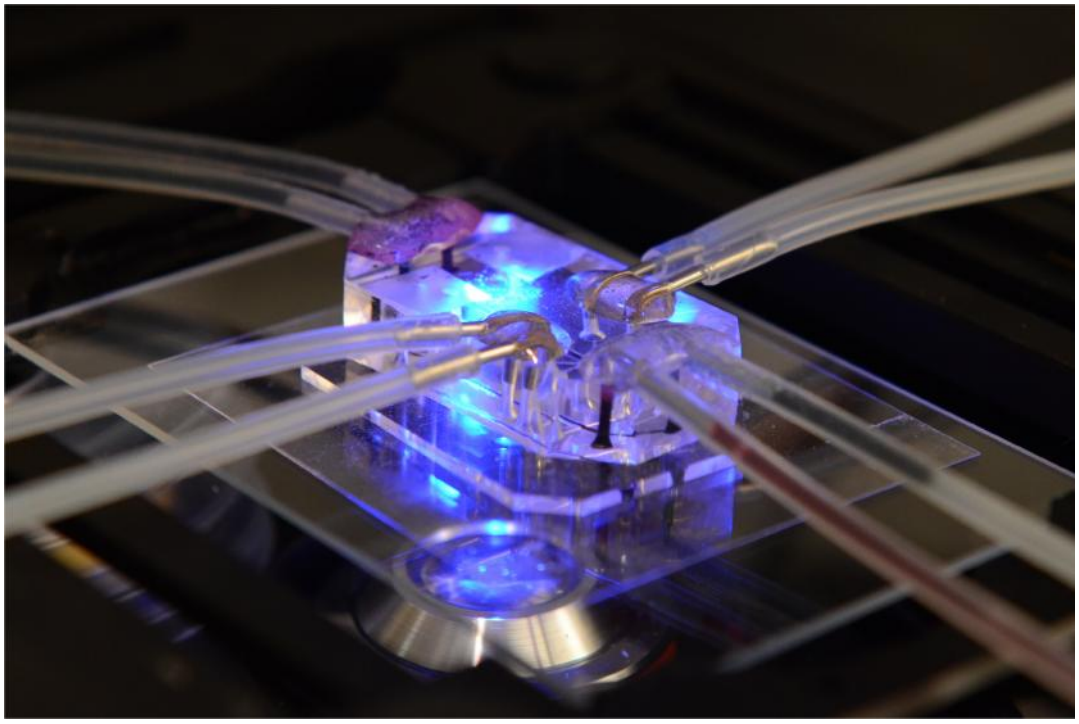

Fig. 1: Lung-on-a-chip on a microscope
This microphysiological system recreates the alveolar-capillary interface, one of the major functional units in the lung and the site where oxygen enters the body. This same interface is where aerosol-based drugs are delivered, where some cancers can metastasize, and is a major site where pneumonias develop, among other things. The lung-on-a-chip is the size of a computer memory stick.

The Wyss researchers have also successfully mimicked pulmonary edema, a deadly condition in which the lungs fill with fluid and blood clots, using the lung-on-a-chip (Huh et al., 2012). In experiments using the cancer chemotherapy drug interleukin-2, or IL-2, the researchers showed that it caused pulmonary edema in the lung-on-a-chip, just as it can in humans. When they injected the IL-2 into the lung-on-a-chip's blood channel, fluid leaked across the membrane, reducing the volume of air in the other channel and compromising oxygen transport. What surprised the researchers was learning that the physical act of breathing greatly enhanced the effects of IL-2 on pulmonary edema. Finally, collaborating with the pharmaceutical company GlaxoSmithKline, they were able to identify a drug that prevented pulmonary edema in the lung-on-a-chip, and this same drug was shown by the company to also prevent pulmonary edema in animal models. These results showed that the lung-on-a-chip could be useful for developing a human disease model as well as for assessing drug toxicity and efficacy, Ingber said.

Since then, Ingber and his fellow researchers have had similar successes with recreating other organs on chips.

Feedback provided by the FDA suggested that to prove useful for toxicity testing, a new device such as the lung-on-a-chip "simply needs to be as good as or better than an animal model," Ingber said. He feels that devices such as the ones his lab are developing are good candidates because they are at least as good as animal models, but they utilize human tissue, and the results obtained so far are robust and statistically significant. He said that the Wyss Institute is presently trying to validate their organs-on-a-chip by working with pharmaceutical partners that are willing to test new drugs using both the new device and animals to generate comparative data. To advance the lung-on-a-chip so that FDA might accept it as an alternative for toxicity testing is challenging, he said, because if a pharmaceutical company has already spent time and money on animal testing, they are not likely to want to invest additional funds to validate new devices or tests, no matter how promising they are.

To Ingber, this underscores the value of getting regulators involved in the process of validating these sorts of systems. "I think we want to think about a new model for validation," he said. He suggested that regulatory agencies may want to consider validating such new devices and tests for their own internal use.

2.2 Industry perspective

The next speaker in the first session was James L. Stevens, Ph.D., who discussed microphysiological systems from an industry perspective. Stevens is a distinguished research fellow at Lilly Research Laboratories, the research and development arm of Eli Lilly and Company, one of the world's largest pharmaceutical corporations.

Stevens began by providing his industry's perspective on the current state of affairs regarding the value of medicine and the effectiveness of preclinical safety testing. "The belief system within the industry is that medicines do in fact reduce morbidity and mortality and improve patient outcomes," Stevens said. "All medicines have a risk/benefit ratio. There's no

medicine with zero risk. We hope the benefits greatly outweigh the risks." He continued: "Reducing or eliminating the use of animal testing is necessary and desirable, but it should not increase risk to humans." He also stressed that, in order to deliver innovative new medicines, a sustainable scientific, regulatory and business environment is required.

Next, Stevens addressed the role played by preclinical safety testing. He pointed out that the focus of preclinical testing is safety in clinical trials, not post-marketing safety. Therefore, the preclinical testing done using animals is aimed at assessing safety for the clinical trials. In turn, the use of human subjects in clinical trials is intended to predict a new drug's safety in humans. "In this context, safety could be defined as an absence of toxicity within the clinical exposure range," Stevens said.

Stevens went on to analyze why drugs fail in clinical trials. There's a perception that the way we go about assessing the risk of new medicines is not working well and does not necessarily protect patient safety well, Stevens said. This is based on the much-publicized analysis by Kola and Landis published in *Nature Reviews Drug Discovery* in 2004 showing that pharmacological toxicology studies predict toxicity in humans only 70% of the time (Kola and Landis, 2004). This study also shows that 14% of all pharmaceutical compounds that fail do so because of a lack of clinical safety.

An analysis by Lilly Research Laboratories frames the same findings in a different light, Stevens told the audience. The Lilly analysis shows that because attrition due to all causes in clinical development is 90%, 90% of 14% – or 12.6% – of all compounds entering clinical development are likely not to reach market due to clinical toxicity. "It's a minor adjustment, but it can be important," he said.

Stevens said that the best dataset the industry has for suggesting how well pharmaceutical toxicology testing works is the Olson data (Olson et al., 2000). He showed slides explaining how the "Olson trilogy" evaluates all occurrences of adverse events known to happen in humans in the context of clinical trials. The data shows that if you go back and look at preclinical testing safety, 71% of the time there was a signal in the preclinical data, he said. "The data are different for different organs, but out of that 71%, the signal was present in 94% of the 30-day studies that enable the first human tests. Putting those numbers together suggests that animal tests uncovered a signal 67% of the time.

This raises the question about whether animal testing should be used to predict target-organ toxicity or the absence of toxicity (or safety) in clinical trials. "Predicting target organ toxicity means we predict the nature of the adverse event that classifies a molecule as unsafe," he said. On the other hand, if the purpose of animal testing is the absence of toxicity, he argued that safety can be considered the absence of unmanageable toxicity within the clinical exposure range. "Our goal in the pharmaceutical industry is to conduct a clinical trial where there is an absence of human toxicity…. not necessarily to predict what's going to happen if" there is toxicity, he said.

Stevens used what he termed a "confusion matrix" to walk through the Olson data by breaking it into true positives, false positives, true negatives and false negatives. He contended that the data shows that the pharmaceutical industry can theoretically achieve what he called 96% "negative predictive value" to predict when there are no adverse outcomes associated with toxicity. This is based on animal studies at exposures covering the predicted clinical exposure.

Using this strategy can reduce the numbers of animals used by advancing better compounds that are more likely to test negative for adverse effects, he explained. To get there, the industry needs "to use the data we have more effectively in our decisions," he said.

Next, Stevens talked about his thoughts on how microphysiological systems can help improve safety assessment by avoiding toxicity, predicting target organ toxicity and managing risks in clinical trials. He used an information-packed build slide titled "Results vs. Compounds" to show where *in silico*, *in vitro*, and *in vivo* data was most useful in the stages of drug development. In each stage of the process, "I'm making decisions based on different types of information," Stevens explained.

At the beginning, it is important to know a lot about the libraries of chemicals that are similar to the one being developed, so *in silico* data can be very useful for making predictions and quickly generating data for all of the similar molecules. *In vitro* tests can be useful for scaffold selection and lead optimization. As the process moves forward, *in vitro* tests can be used to predict *in vivo* responses.

Towards the end of the process, when developers have hundreds of data points about some similar molecules, they can use all three kinds of data to good end. By the final stage, when researchers have converged on a single candidate molecule, *in vivo* tests are most important.

"If I can avoid risks to improve the probability of negative *in vivo* outcomes, I'm going to reduce the number of *in vivo* studies if I have to", Stevens said. This includes "anything I can do to manage the risk for humans, to frame that risk so that it is manageable in the context of clinical trials," he said.

Stevens said that Lilly tends to use microphysiological systems "on the back side", i.e., after a negative outcome has been brought to light, to replicate the pharmacology and biology to see if the problems can be avoided or managed. For example, Lilly scientists have used human cardiomyocyte heart muscle cells derived from induced pluripotent stem (iPS) cells to identify pharmacology-based impacts on cardiovascular function. He said that when pharmacologists know a target organ, microphysiological systems can also focus screening. Areas for improvement include developing microphysiological systems that evaluate and/or model the biological response networks relevant to the target organ physiology and pathophysiology, he said.

2.3 Regulatory perspective

The final speaker in the first session was Douglas Throckmorton, M.D., deputy director for regulatory programs for the Food and Drug Administration's (FDA) Center for Drug Evaluation and Research. He discussed microphysiological systems from a regulatory perspective.

Throckmorton's job requires him to identify things that the center needs to do but is not presently doing. The center is very interested in microphysiological systems in terms of their use in the development of new products for the U.S. public, he said. The FDA has processes in place that can help facilitate the use of these new tools.

New drug success rates are not as good as they could be, Throckmorton said. He pointed to articles showing that the success rate of new drugs in phase II trials is 18% and the rate in phase III trials is 50%. The 50% phase III trial success rate is particularly problematic, given that companies have invested hundreds of millions of dollars by that point and hundreds to thousands of patients are exposed (Arrowsmith, 2011). There are lots of reasons why drugs fail throughout the medical pipeline, but Throckmorton said that his discussions with experts at some of the nation's largest drug companies revealed that they all had different explanations for the failures. Major sources of failures include lack of efficacy, toxicity, or issues related to the potential drug's absorption, distribution, metabolism and/or elimination. The challenges that drug developers face include the heterogeneity of disease related to differences between animals and humans and *in vitro*. The American public expects drugs to be safe, Throckmorton said.

To address these problems, FDA needs to support the development of predictive physiological biomarkers, Throckmorton continued. The process through which biomarkers are currently generated, which tends to be on a case-by-case basis driven by drug manufacturers' needs with a slow movement towards general use as scientific experience accumulates, is inefficient.

In this context, Throckmorton perceives that some of the new tools being developed, including microphysiological systems, offer "profound opportunities." In addition to being non-animal-based, they can be highly efficient in terms of the number of products that can be screened. They also have the potential to be integrative, in that they offer the potential to answer more than one question about a drug's effects.

The challenges to using microphysiological systems include the fact that in some ways they are device-like, which requires additional regulatory scrutiny. Their utility for drug development may also increase the regulatory complexity associated with their use. Both the novelty and sophistication of these systems may also hinder their use, he added. "This is a common challenge for new technologies and science," he explained.

The FDA's efforts to foster the use of new technologies such as microphysiological systems include its Voluntary Exploratory Data Submission process meetings. The FDA holds these meetings with groups working with new science, and the information is exchanged with the promise of no regulatory impact. The goal is to provide academic and industry groups insights into the FDA's scientific perspective on new technologies in a given context. The meetings also provide a mechanism for FDA scientists to learn about new science in advance of regulatory submissions. The agency has had more than 50 of these meetings since 2004, and it has increased the numbers of regulatory submissions with novel biomarker data. It has also helped the FDA with policy development.

FDA's Drug Development Tool's Qualification Program also provides a framework for speeding regulatory acceptance of scientific tools for use in its drug development program. The program currently focuses on biomarkers, clinical outcome assessments, and animal models for use under the agency's Animal Rule.

Scientists developing or refining biomarkers for a specific context of use can participate in the Biomarker Qualification Program. The data generated will support rigorous evaluation of the submission for use in the regulatory process. Once FDA has qualified a biomarker, the sponsors can use it in the qualified context without requiring the agency to reconsider and confirm it after it has been fully developed. Thus far, the agency has qualified about 20% of the products that have come through. The process has facilitated the development of biomarkers because it results in binding regulatory decisions.

Approximately 43% of the biomarkers that have been submitted through the Biomarker Qualification Program relate to preclinical and clinical safety.

From the FDA's perspective, the more focused the approach used to develop microphysiological systems is, the easier it is to qualify. The developer's approach to qualification needs to be rigorous and systematic, Throckmorton stressed. "Don't choose the data that support your system and ignore the data that don't support your system," he cautioned. "We're going to demand to look at all of the data."

Throckmorton concluded by telling the audience he was confident that the FDA's programs will help foster the development of microphysiological systems.

3 Microphysiological system cell types, tissues and applications

This session was chaired by Thomas Hartung, the Doerenkamp-Zbinden Professor and Chair for Evidence-based Toxicology at Johns Hopkins University, as well as the director for the Center for Alternatives to Animal Testing (CAAT). Hartung also gave the first talk on the topic of Good Cell Culture Practices and Quality Control.

3.1 Good Cell Culture Practices and Quality Control

Over the past two decades, Hartung has been very interested in the quality of *in vitro* systems. His focus on cell culture practices dates to 1996. Between 2002 and 2008, he headed the European Centre for the Validation of Alternative Methods (ECVAM). More recently, he has been involved in efforts to determine how to validate methods for evaluating medical countermeasures and drugs to bioterrorism and warfare. The respective National Academies of Science panel he served on concluded that there is no animal model for testing countermeasures to bioterrorism. His recommendation is to use alternative methods to animal models

that are a marriage of cell culture and bioengineering. CAAT is now involved in designing quality assurance and validation assessment approaches for the inter-governmental programs that are developing these countermeasures.

Like earlier speakers, Hartung pointed to the high percentage of failures of new drugs by the time they reach clinical trials. He enumerated some of the limitations of animal-based drug development beyond the reality that humans are not 70-kg rats. These include the use of young animals, artificial diseases, unrealistic treatments, and the lack of co-variables such as comorbidities and other treatments. Another concern is research suggesting that only between 11% and 25% of academic studies are reproducible when analyzed in two recent prominent articles by industry scientists.

Further, there have been relatively few systematic evaluations of how well animal models mimic common human diseases (Pound et al., 2004; Hackam and Redelmeier, 2006; Rice, 2012). The results of the evaluations that have been performed on stroke, sepsis, and multiple sclerosis have been disappointing. Hartung pointed out that a critical review found that genomic responses in mouse models do not do a good job of mimicking human inflammatory diseases (Seok et al., 2013).

The challenge starts with cell culture, Hartung said. He gave an example of this from around a decade ago when his lab obtained commercially available genetically engineered cell cultures that were transfected with cytochrome P450. When his group karyotyped the cells to identify the chromosomes, they discovered that in some cases the chromosomes were incomplete or contained additional fusions of chromosomes. Cells that were definitely not the same genetically would at the time have been interpreted as being different in only one additional P450, he said.

Additional research showed that many cells are neither the cell type or even from the species they are purported to represent, Hartung said. About a decade ago an analysis showed that 10-11% of cells in cell banks are not the species they are supposed to represent.

In Hartung's analysis, the human-on-a-chip microphysiology approach can help overcome important shortcomings associated with using *in vitro* cells. The fact that they have the potential to mimic the differentiation of organs is a positive, as is their potential to supply oxygen to the cultured cells. Another plus is that microphysiological systems are likely to overcome problems associated with the tumor origin of many cell lines as they mainly make use of iPS cells.

However, scientists do not yet have a plan to stress the human-on-a-chip cells continuously, so they are likely to be as "bored" as conventional *in vitro* cells, which Hartung sees as a key reason for dedifferentiation *in vitro*. There is as yet no way to mimic metabolism or stimulate defense, either. Thus far there also are no plans for introducing analytics to determine the fate of test compounds in culture.

3-dimensional (3-D) cultures offer some improvements over 2-D cultures, Hartung continued. These include increased cell survival and differentiation. The 3-D cultures also allow for increased cell-to-cell interaction, and they do a better job of reproducing the complexity of human organs.

However, the endpoints still need optimization, Hartung said. The main issue is that more complexity means lower reproducibility and throughput. "We have to do a lot of compromise with these 3-D systems," he summarized.

Although 3-D models based on human cells cannot help overcome all of the shortcomings currently facing toxicology, they can target some of them, Hartung said. These include reducing the use of test animals, expanding scientists' ability to evaluate new products and new hazards, and the ability to give some indication of individual variability. Testing with 3-D human cells may also prove to increase the predictivity of toxicology testing. But at present it appears not to have too much of an effect on the high cost and slow throughput of classical toxicology, and it may be too precautionary.

Hartung contended that the process of validating some of the alternative toxicology tests being developed is one of the rare examples of quality assurance in biomedical research. "Only by proving quality can we move ahead," he said. His group is pursuing the idea of using the principles established in evidence-based medicine to produce evidence-based toxicology (Evidence-based Toxicology Collaboration[2]).

Quality assurance also encompasses good laboratory practices. Good laboratory practices have been established for animal-based tests used in toxicology, but the toxicology laboratory practices can be very poor for non-animal-based testing, Hartung said. A drawback to the established good laboratory practices is that they often cannot be implemented in academia due to their costs and lack of flexibility.

The OECD has produced a guidance document to establish good laboratory practices for *in vitro* tests, and other groups, some in collaboration with the OECD, have worked to establish good cell culture practices (GCCP) (Coecke et al., 2005) and good validation practices. Hartung has been a key player in the effort to establish GCCP, which deal with the inherent variation of *in vitro* test systems compared to *in vivo*. While an animal maintains homeostatic conditions, the conditions for *in vitro* cell culture are by definition artificial, he explained. The aim of GCCP is to reduce uncertainty in the development and application of *in vitro* procedures by encouraging the establishment of principles for the greater international harmonization, rationalization, and standardization of laboratory practices, nomenclatures, quality control systems, safety procedures and reporting. The documents and guidance that have been produced are intended to support best practices in all aspects of the use of *in vitro* systems, including the use of cells and tissues. Where appropriate, these principles are linked to the application of good laboratory practices.

Validating organs-on-a-chip presents a new challenge, Hartung told the audience. There is a dramatic difference between a model and a test, he explained. Just as liver cells are a model, microphysiological systems are models, not tests, he continued. A given test is defined by its purpose, Hartung stated. It is defined by a very precise protocol. For example, hepatocytes can be cultured in different ways to produce an endless number of different tests, depending on the intended goal and how the parameters are set up.

[2] http://www.ebtox.com

While animal models have been mostly accepted based on their "face validity" because they use a healthy living organism, microphysiological systems are essentially devices with a variety of different elements. As models, they must be defined by precise protocols based on the phenomenological similarity to the organisms and/or current scientific understanding.

From there, tests must be defined by their purpose and the approach for their mechanistic basis. This needs to include the case for their relevance and a protocol that provides a precise description of how to read the experiment. The protocol information needs to include details about the standard operating procedures, a specification of endpoints and endpoint measurements, and how they are to be interpreted. Adequate negative and positive controls must also be included, as well as an indication of the test's limitations.

In this context, validation is the independent assessment of a test's scientific basis, reproducibility and predictive capacity, Hartung said. Historically, new alternative tests used traditional animal tests as points of reference, which does not necessarily address their scientific basis, and focused on their predictive capacity only. Scientists are increasingly faced with the challenge of validating new tests that do not reproduce animal tests. Examples of this are tests based on human physiology and ones based on toxicity pathways.

This brings to light the need for different reference points for validating new studies. Expert consensus is necessary, but in the end it is important that we identify a composite type of reference points that replicate scientific knowledge.

Hartung reiterated his belief in the practice of evidence-based toxicology as a way to assure the quality of and to validate new toxicology tests. Hartung first articulated his vision for how the principles established in the field of evidence-based medicine could be applied to toxicology in 2005 and developed the concept further the following year (Hoffmann and Hartung, 2006). The first international forum on evidence-based toxicology was held in 2007 and the Evidence-based Toxicology Collaboration was launched in the U.S. in 2011 and in the EU in 2012. Both the US and EU collaborations have distinguished steering groups that are working to establish how the tools of evidence-based medicine can be applied to toxicology. These objective tools include systematic reviews and meta-analyses.

Evidence-based toxicology's approach to validation does not require reference tests. Instead it focuses on mechanistic validation using mechanism as a substitute for correlative predictivity, Hartung explained. The idea is to establish a reference based on an established mechanism, the Human Toxome (The Human Toxome Project[3]). The scientific knowledge base on the mechanism is used as a point of reference. This generates a proof that a given pathway is covered. The process also establishes reproducibility.

Hartung concluded by pointing out that although toxicology is serving as the pilot for validating how evidence-based approaches can be applied, all areas of the life sciences have similar needs. He stressed that the validation process is constantly evolving, and the principles for validation defined by the EU-based ECVAM and taken up by the US-based ICCVAM and internationally by the OECD have been shown to work. Finally, he repeated his belief that evidence-based toxicology is an important tool for validating 21st century toxicology methods.

3.2 The importance of stem cells

The second talk in the session was by Kyle Kolaja of Cellular Dynamics International (CDI), a company that manufactures human induced pluripotent stem cell (iPSC)-derived tissues under the iCell® and MyCell® brands. He made a case for why stem cell-derived tissues can help researchers develop microphysiological systems.

Stem cell-derived cells are likely the most ideal format available for generating the cells and tissues used in microphysiological systems, Kolaja said. Quality management is important for manufacturing a reproducible and robust human cell product, and this same level of rigor is required to produce reproducible and robust microphysiological systems, he stressed.

Kolaja reminded the audience that the world of cell-based biology was turned on its head in 2006 when a team led by Shinya Yamanaka of Kyoto University in Japan published research describing his success in inducing somatic mouse cells to become stem cells capable of producing a variety of cell types (Takahashi and Yamanaka, 2006). In 2007, both Yamanaka's group and a team led by James Thomson at the University of Wisconsin at Madison reported a method for converting human skin cells into cells that very closely resemble human embryonic stem cells (Takahashi et al., 2007; Yu et al., 2007). "This discovery changed the way we look at cell fate and cell biology; to reprogram an adult cell back into a state where it is renewed in perpetuity and can differentiate into all of the different tissues in the body reflects a fundamental shift in cell culture," Kolaja summarized.

Kolaja stressed that the process of producing terminally differentiated tissues from induced pluripotent stem cells is not as easy as it appears. "When it is done correctly, the cells will be produced in ample supply, high purity and high quality to do lots of interesting experiments." Tissue-derived iPS cells are crucial for producing the complex 3-D microphysiological systems that have multiple cell types communicating with each other, he says.

Cell culture itself dates back to a little over 100 years ago to when Ross Harrison of Johns Hopkins University first cultured frog neurons (Harrison, 1910). Almost 60 years ago, the first immortalized human cell line (HeLa) was derived. Since these fundamental first steps, immeasurable innovations and advances have brought cell culture to its current state.

From the perspective of toxicology, the limitations of primary cell culture have hindered the potential of replacing animal and human experiments, Kolaja continued. The challenges associated with making the shift include consistent access to primary human and animal cells, variability due to how cells can be isolated, and the degeneration of the "*in vivo*" phenotype once in culture. By addressing these issues, stem cell-derived tissues can also help further the development of the 3-D models needed to produce microphysiological systems.

[3] http://humantoxome.com

A key advantage of stem cells results from how they can be derived via genetic engineering, Kolaja told the audience. “The iPSC field has moved past the early days of integrated viruses,” he said. Now methods that do not require integration in the genome are used predominantly to reprogram cells. Small amounts of peripheral blood or other tissue can be used as starting material. The iPS cells are grown in defined media and on well characterized matrices, two improvements that have helped provide consistency to stem cell culture,” he explained.

It is also now possible to insert genetic selection markers in order to purify a specific cell type of interest, Kolaja said. Transcriptional factors can also be engineered into a cell to help drive differentiation down specific cellular lineages, thus increasing efficiency. These methods leverage modern molecular biology knowledge and genetic engineering approaches, with new techniques being developed every day, he said. The state of the art of genetic engineering can be brought to bear on iPS cells to develop cell types of interest, induce genetic disease models, and better reproduce human cellular biology. This is one of the major reasons why, from a manufacturing perspective, iPSC-derived cells give scientists a tremendous advantage over primary cultures.

Unlike primary cultured cells, which often dedifferentiate or lose their phenotype in culture, stem cell-derived tissues demonstrate a maturing phenotype *in vitro*. In fact, iPSC-derived cells, such as iCell® Hepatocytes (Fig. 2), tend to mature over time. This allows them to be incubated over extended periods while continuing to accumulate an adult phenotype.

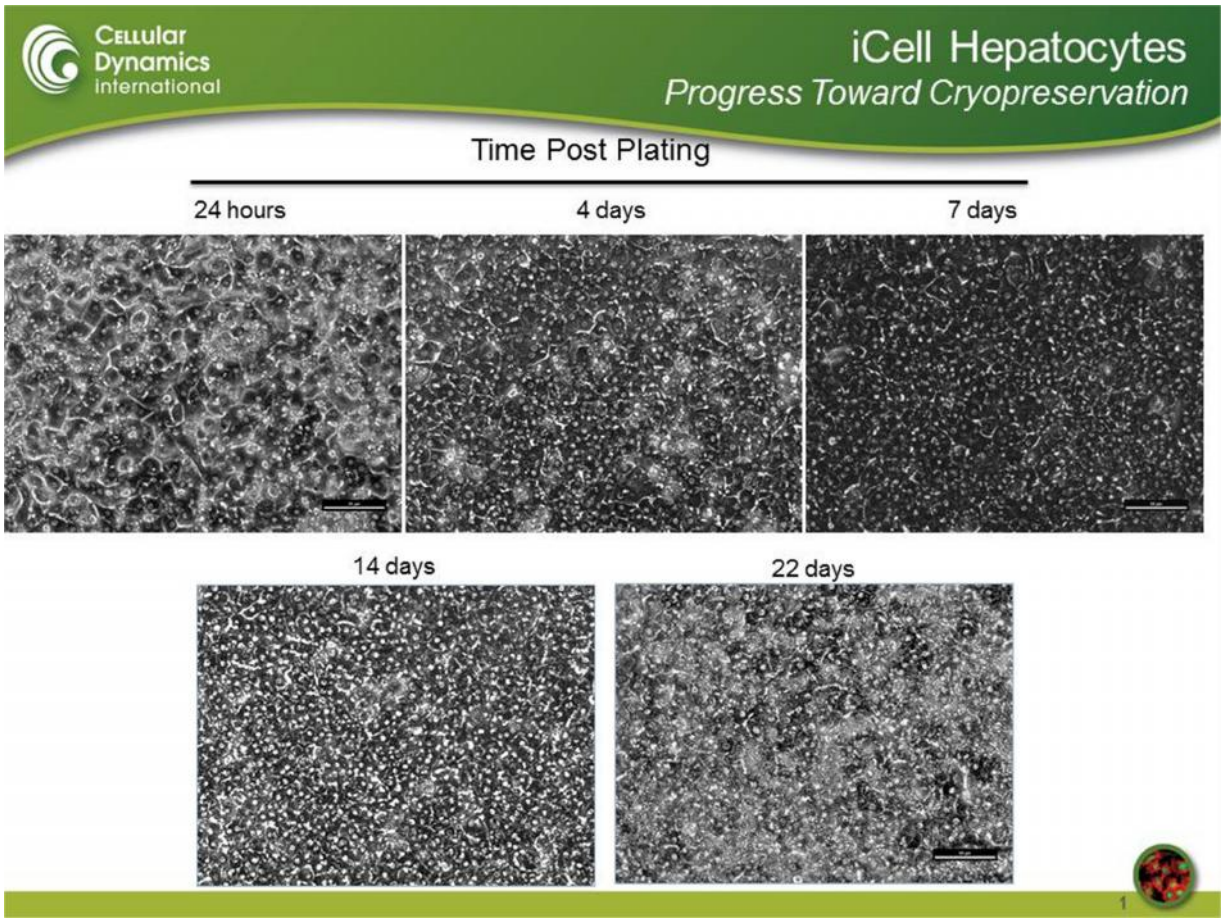

Fig. 2: Maturation of iCell® Hepatocytes
Like other iPSC-derived cells, iCell® Hepatocytes tend to mature over time. This allows them to be incubated over extended periods while continuing to accumulate an adult phenotype.

“Scientists are actively trying to find ways to speed up the maturation of stem cell derived tissues, as the fetal expression pattern of stem cell-derived tissues can be considered a deficiency,” Kolaja commented. Gene expression research shows that iCell® Cardiomyocytes are more similar to adult human heart samples than primary cultures (Babiarz et al., 2012). Research also demonstrates that subjecting them to electrical stimulation can speed their maturation. Small molecules additionally show promise for helping enhance the maturation of iPS cells. Other 3D culture methods, such as bioreactors, have shown promise for maturing stem cell-derived tissues. Kolaja also pointed out that adult phenotypes are not always needed for toxicological research.

iCell® Cardiomyocytes express major cardiac proteins, and they respond to drugs such as channel blockers as expected. They look essentially like cardiac muscle cells and have functionally active cardiac channels. They also exhibit cardiac action potential wave forms. Researchers have used iCell® Cardiomyocytes to evaluate how different drugs can impact cardiac rhythms. Scientists have also discovered ways to induce the iCell® Cardiomyocytes to thicken to emulate the hypertrophy that can cause sudden cardiac death.

Kolaja also told the audience that iCell® Hepatocytes can be infected with viruses to study infectious diseases such as viral hepatitis. He pointed out that iCell® products are delivered as cryopreserved cells, and regain their functionality once they are thawed.

The last piece of why iPS cells hold such potential is because of their ability to allow researchers to control genetic diversity. This opens the possibility of allowing scientists to take samples from anyone. It allows researchers to recapitulate disease phenotypes, and it opens the door to prospective and retrospective clinical trials. This feature also enables scientists to try to get a handle on phenotypic diversity as it relates to *in vitro* models.

Many labs are currently attempting to produce tissues derived from iPS cells, Kolaja said. The challenges to manufacturing these cells in a robust, reproducible manner parallel the requirements for drug development in some ways. The cells that are produced must be of high purity, and they need to be generated in large quantities to be able to study them further. Cryopreservation is a key component to "stockpile" iPSC-derived cells for distribution around the world.

Quality management is crucial to manufacturing consistent batches of iPS cells, Kolaja pointed out. Every single component used in the process, including the reagents, needs to be tightly managed with an enterprise-wide approach, he said. This includes documenting all changes and keeping extensive records to track all aspects of manufacturing and to gain insight into changes discovered in the cell lines.

Cellular Dynamics is now working to create biobanks of iPSC lines and terminally differentiated tissue cells. For example, the company is collaborating with the National Institutes of Health's Heart, Lung, and Blood Institute in a project focused on left ventricular hypertrophy that involves patients from a larger genome wide association study. The project involves deriving iPS cells and then making cardiomyocytes from the most informative 250 individuals, including those with the hypertrophy phenotype, and performing functional analysis. Ultimately, the researchers hope to correlate the *in vitro* findings with results from genome-wide association studies aimed at confirming the functional deficits of common genetic factors that influence development of the disease.

Cellular Dynamics has also been awarded a grant from the California Institute for Regenerative Medicine involving 3000 patients with a variety of different diseases, including Alzheimer's, autism spectrum disorders, autoimmune diseases, cerebral palsy and diabetes. Cellular Dynamics will be making 9000 iPS cell lines from these individuals and, as a subcontractor to Coriell Institute for Medical Research, will be biobanking them. Researchers can work with the banked cells themselves or contract with CDI to differentiate the cells into specific cell types of interest.

Kolaja concluded by restating his belief that the opportunities for stem cell-derived tissues are tremendous. He believes that it will be possible to develop an *in vitro* model using iPS cells that will ultimately supplant animal studies.

4 Panel discussion

The moderator for the panel discussion, Susan Fitzpatrick, senior advisor for toxicology at the FDA's Center for Food Safety and Applied Nutrition, began the session by asking whether any of the panelists in the first two sessions had questions for each other.

Thomas Hartung said that he had hoped to hear more about where "we are not yet." The speakers who had addressed the conference were very optimistic about the potential of these new systems to solve problems. But he pointed out that experience shows that there are difficulties associated with getting enough cells to use in the systems. Our problems include the incomplete differentiation of the iPS cells and the lack of harmonized protocols. "I think there's a lot still to do and I hoped that you could add a little about these components," he said.

Kolaja responded that many of the challenges are due the fact that iPS cells are made in small scale, they are produced in ways that are challenging to reproduce, and the approaches to producing them result in relatively less pure populations of cells. "It is difficult to have a reproducible, robust system. It takes a long time to get where we have a locked-down system." The manufacturing challenges include the selection of media and reagents. The challenges around the maturation of the stem cells are fundamental, he acknowledged. Cardiomyocyte heart muscle cells are an example of this, he said, noting that Cellular Dynamics has done a lot of work with arrhythmias. Even though the stem-cell derived cardiomyocyte cells have a fetal phenotype, they still work fairly well to predict human responses.

On the other hand, hepatocyte liver cells have a mixed phenotype, with elements of fetal and more developed cells. In order for iPS-derived hepatocytes to be functional, scientists need to find a way to mature the cells. "We're in the early days of learning how to use this technology. A lot of our challenge as a company is in helping people change their paradigm from being used to getting things with rodent primary culture, or human primary culture, which have fundamentally different techniques." He reiterated his belief that the potential of iPS cells justifies the effort required to find solutions to these challenges.

Next, David Jacobsen-Kram, associate director of pharmacology and toxicology at FDA's Center for Drug Evaluation and Research's (CDER's) Office of New Drugs, asked Kolaja whether the various differentiated iPS cell types will be able to grow in the same culture medium that is going to circulate among them. Jacobsen-Kram also asked Stevens if it was correct to impute that in the more acute clinical trials for safety a 30-day study is sufficient.

Kolaja said that he attended some of the early sessions when Defense Advanced Research Projects Agency (DARPA) and National Institutes of Health (NIH) were first proposing the idea of producing a human-on-a-chip. He said that he initially expressed concern about finding a medium that will enable a lot of different cell types to communicate and "be happy." The program has a lot of emphasis on engineering and flow-through systems, and he believes it is possible for each organ chip to have a very generic medium that it is able to move through. The organ-specific micronutrients that are crucial to maintain that cell community could be introduced at the beginning of each flow-through part. "I don't think it's impossible, but I think it's going to be a critical engineering component," he said.

Ingber added that one of the major challenges was coming up with a universal blood substitute. He noted that he has one medium in his body that keeps all of his cells alive. "We flow blood through these devices in other projects all of the time," he said. His team is currently exploring if their epithelium lined vessels can use blood and plasma. That said, he agreed that it will be challenging to optimize media depending upon organ type. "If it is possible to use plasma and blood it will be helpful for issues like plasma-XIII binding," he said.

Stevens responded to Jacobsen-Kram's second question. "Sufficient is a very subjective characterization. Is it ever sufficient if we do encounter adverse events in the context of clinical trials? However, from a pure assay metrics point of view, I think what it says is, when we get it right, a 30-day study gets it right in producing the signal 90% of the time. He noted that the numbers he showed were from phase I and II. It is important to look at adverse event rates from phases I, II, and III. "The reason why we're doing it this way is because that's the way the Olson data was constructed," he explained. He added: "We haven't actually looked at the predictivity or the sensitivity or specificity by days matched to the length of the toxicology study. So it is possible that some of those later-stage failures are actually due to signals that we may have seen in 30-day studies or in longer-term studies. The answer is we just don't know."

Ingber pointed out that the main reason he focused his comments on toxicity testing is because he was at the FDA. All of the pharmaceutical companies his group has met with are equally interested in using microphysiological systems for disease models and efficacy testing. "If you can do all of this with the same system and you can see efficacy while looking at toxicity, potentially with blood flowing, there's a lot of problems we can solve," he said.

Stevens agreed. "What you're proposing is that we can actually tackle this false positive rate. Can we actually say that we now understand the potential benefits much better and can we run a manual clinical trial to see if the benefits warrant the risks?"

Susan McCune, deputy director of FDA's Office of Translational Science, who joined the conference after Throckmorton left to attend another meeting, commented that FDA's animal model qualification program is under the animal rule for medical countermeasures development and does not include the microphysiological systems discussed at this conference. She said that she thought Throckmorton focused on the drug failure rates because he was trying to make a point about safety and efficacy barriers. "A better understanding of those populations where we might see responses and mitigate safety issues is critically important. So we're looking at all of these innovative types of designs, including innovative trial designs and innovative approaches to looking at these populations to mitigate some of these failures from both a safety and an efficacy perspective."

An online questioner asked whether microphysiological systems could model diseased organs.

Ingber responded that in his second talk later in the afternoon he would discuss how the systems were being used to model diseased organs. "For example, we're getting primary cells from asthmatics and people with COPD," he said. "I think that the goal is to be able to model diseased organs as well as normal healthy ones."

Hartung added the beauty of the iPS cells is that the researchers know the genetic background of the donors. In one of his projects, researchers are trying to use neurons derived from iPS cells from children with neurodevelopmental disorders. "This is really the prospect of getting for the first time human variability into the systems," he said.

Lewis Kinter, Senior Director at AstraZeneca, observed that when he began his career in the early 1980s his job was to conduct pharmacological bioassays. He used animal models to look for drugs that would lower blood pressure. There was no thought as to what the pharmacological mechanism might be – it did not matter. The only issue was: Did the bioassay produce the activity that the scientists were looking for. But the industry rapidly evolved toward using a target-directed molecular model. Today very few of those pharmacological bioassays are still used, except perhaps in the realm of safety pharmacology, he said. However, pre-clinical toxicology is still fundamentally a bioassay-driven area, Kinter continued. "We put compounds in the animals, and we look for adverse effects. And you can argue that the clinical program is a human bioassay. What seems to have not happened in the safety assessment realm is much interest in looking at the fundamental mechanisms that are responsible for the small-molecule and large-molecule-induced organ toxicities." He suggested that Stevens' presentation argued for using one preclinical bioassay or sets of preclinical bioassays to predict the results in another bioassay. "Whether you're doing *in vitro* prediction of what's going to happen in *in vivo* animal bioassays, or some more Tox21-type effort involving combinations of bioassays that will predict a negative in a human bioassay, I would challenge you, if you've learned anything on the pharmacological side, why not focus on the fundamental mechanisms and get out of the bioassay business," he asked. Kinter posited to the group that he believed "the fundamental benefit of the 3-D microphysiological systems will not be in their use as yet another bioassay, but will be as a tool to help us get to fundamental mechanisms much more rapidly than we could if we had to use the whole-animal platform."

Stevens responded that he believes Lilly is doing exactly what Kinter is proposing "but in a slightly different way." Stevens said that he spent a lot of his academic career studying reactive intermediates and trying to define mechanisms. "I'm actually not that concerned about the mechanism of a highly reactive species that interacts with a whole variety of cellular targets," he said. Instead, his company is trying to eliminate what he terms "compound property related intrinsic toxicity," rather than to fundamentally understand why they cause cell death.

It can be possible to eliminate possible effects by tweaking the basic functional groups that are not required for a molecule's potency, Stevens continued. "Now when I get into my animal studies, I am faced with what I'm going to conclude is likely to be a pharmacologically based event." For that reason, he said, he does need to understand the mechanism and has a higher likelihood of identifying the mechanism using the tools that are now available. "So we're actually trying to separate the two problems. In the latter case, we're doing exactly what you're proposing in an integrated physiological system – and if we can replicate that which we want to avoid or manage in a microphysiological system, we can screen to a better compound."

Kinter responded that his concern is in Stevens' cell death assay. "You don't know whether the mechanism that was responsible for the cell death was actually a purely toxic mechanism or whether it was a superpharmacological expression that might actually be useful," he said.

Stevens said he realizes "that the predictive system that we're proposing to simplify systems isn't perfect, but – either way – if I have other choices that I say the probability is lower, then I will have something that gives me a good *in vivo* preclinical profile, why would I want to pause and sort out the mechanisms when I can actually go in a different direction?" He said that when he is working with a compound where there is a suggestion of excessive risk based on the current paradigm of using preclinical data to describe the safe operating parameters in a clinical trial, he needs to describe the mechanism so he can do an effective extrapolation from preclinical physiology to human physiology to manage risk. "So I think we're doing it, but I don't see the need to actually understand the mechanism behind a probability-based analysis that suggests I have a low probability of being successful regardless of what the mechanism might be," he said. Kinter suggested that they continue their discussion offline.

Ingber picked up on the discussion. "Learning what you don't know" can be important, he began. When his group was working on the pulmonary edema model that he presented information about later in the afternoon, they submitted an article about it to *Science Translation Medicine*. He said, "One reviewer said this should not be published because it's too simplified, there are no immune cells. The other reviewer said 'this is amazing – we just learned that you don't need immune cells for pulmonary edema, which everyone assumes is the case.'" Ingber said that the concern he has as a fundamental scientist is the assumption that we already know enough to do this sort of probability-based work. We do when we're working with an animal, "but you have to think about 30 years ahead, what's coming up." We have no idea what may be happening with nanotherapies, for example, at that time. "The power of the microphysiological system is that you start with the simplest environment that gives you physiological relevance," Ingber said. Then you add to it. Toxicology is moving toward being a knowledge-based discipline, he added.

"I think there are two ways to reason through a problem; you use deductive reasoning or inductive reasoning," Stevens responded. "We're very deductive in the way we interrogate problems. You never know what you don't know. You can turn the unknown unknowns into known unknowns. Uncertainty is very different from risk. But you still need to make decisions if you want to advance molecules. I'm not arguing that one should supersede another. All I'm arguing is that you can use the two separately and effectively, depending upon where you are in the development paradigm," he said. Very early in the development process it is possible to have a choice of scaffolds, each of which contains 100 molecules, and there is an immense number of degrees of freedom. Near the end of the process you may have zero degrees of freedom. "You're moving between two separate paradigms, one of which requires a more mechanistic understanding. The other is a probability-based exercise where it is most important to know that there is a probability of being successful." To Stevens, this approach drives a more efficient business.

Kinter responded by pointing out to audience members not as familiar with pharmaceutical drug development that the place where the investment needs to be made is in the discovery space "long before you even think about queuing up for a GLP study." (GLP stands for Good Laboratory Practice and it refers to safety studies required by regulatory agencies for drugs at the final stage of preclinical development.) He said that both Lily and Astra-Zeneca agree that the moment a company makes the decision to engage in the animal toxicity, genotoxicity and supporting toxicokinetics studies, and safety pharmacology GLP studies required to establish a drug's safety, "drug development basically devolves to a game of high-stakes poker. The trouble is, of the cards that are dealt, some are old and some haven't been revealed yet, but there's nothing you can do to change it. You'll make one of two decisions. You'll either put more money in the pot, or you'll fold your hand. The only thing you can do in poker but you can't do in drug development, is bluff. It's also true that you can be holding a good hand in drug development and still blow it," he said.

Hartung responded by stressing his belief that "mechanism is key." "It's the only way to get us out of the problem of correlation, which is often misleading. It's the only way to get a handle on rare events. You cannot really discover these through statistical methods. In the end, what we look for is rare events, whether it is the rare pharmacological active agent in drug screening or the rare toxic effect in safety testing," he said. "We really need causality to aggregate around mechanism, but a big question mark for me is whether the advantage of the microphysiological systems being closer to real physiology in terms of organ function is sufficient. How does this balance against the cost, complexity and the lower number of replicates I have," he asked. After all, he pointed out, many of the steps involved in addressing mechanism require replicates. It's never good to put all of your proverbial eggs in one basket, he stressed, but microphysiological systems certainly occupy a middle ground between whole body testing and the simple systems that give us the first indications of mechanisms.

James Hickman of the University of Central Florida asked McCune about how these devices are evaluated. There has been some talk about the Center for Devices and Radiological Health (CDRH) having an office on *in vitro* tools, he said. His understanding was that this office was actually for evaluating biomarkers in the clinical phase of testing. In contrast, he understood that the tools being discussed at the meeting are preclinical. Would they be evaluated through the Office of Translational Research? Is there the opportunity to go through both, or does it really depend upon the application, he asked.

McCune said that her office has been in discussion with CDRH regarding how to handle this. “This would be on an individual case-by-case basis,” she said. What the agency is trying to initiate are “opportunities to come in and have conversations with us independent of the traditional drug development program route through the review commissions,” she explained. One of these opportunities is through the Voluntary Exploratory Data Submissions (VXDS) program, the other is through the Drug Development Tool qualification program, she said. “Clearly all of these products that we’re now talking about have components of both associated with them, which will require that we have a collaborative approach.”

Fitzpatrick clarified that the FDA does not perceive a difference between preclinical and clinical. “It’s whether or not the microphysiological system could be used as a device, and not just a drug development tool that could be used in preclinical or safety testing,” she said. McCune agreed.

Melvin Andersen from the Hamner Institutes said that the conference was “giving him a crash course on the FDA.” He asked Kolaja to provide more information about how the cells being used in the microphysiological systems are produced. “A lot of the differentiation *in vivo* is thought to occur when precursor cells move into environments with other partnering cells,” he said. “Do you consider these cells you’ve made to be precursor cells or the cells that would be in organs. How are you looking at those differences,” he asked.

Kolaja replied that the answer depended upon the cell type. “We look at it as the creation of an individual cell type as purified as possible,” he said. One of the aspects that Cellular Dynamics is beginning to consider in making cells such as cardiomyocytes is adding back other cell types related to the heart that enable contractility. Purification is an important component, he said.

Andersen responded by pointing out that Cellular Dynamics is also working with a variety of partners and end users of these technologies. Are these end-users concerned whether subsequent steps of differentiation will require cell-to-cell interactions, he asked.

Of course end users are concerned about this, Kolaja said. “We’re in the early days so a lot of the work that is being done is on single cells and single populations of cells in a monomer,” he said. “As we begin to migrate towards organoids where you have stromal cells, parenchymal cells and immune cells, all of these things will have to be repopulated. We look at it more as starting with a purified population so we can define that mixture rather than manufacturing it all simultaneously,” he said.

Ingber commented that in his experience, scientists create the physical microenvironment as well as the chemical and cellular microenvironments that have “led us to consistently find levels of organization we hadn’t seen before.” He pointed to the CaCo-2 colon cancer cells that everyone uses (but everyone hates), which are a classic tumor-derived epithelial cell line that is not differentiated in culture. “People use it because it’s the best they have,” he said. Ingber said that his group’s effort to use CaCo-2 cells to create gastrointestinal systems-on-a-chip resulted in differentiation into all of the cells of the small intestine, even though the cells came from the colon. “We get basal cell crypts… and recreate the entire differentiation cascade. We have bone marrow-on-a-chip where we use a different approach but we can maintain all of the cells, including hematopoietic stem cells, in the right ratios over time in culture.” To Ingber, this confirms what he knows from biology and medicine that the niche plays a key role – stem cells have stochastic controls that depend upon the proximity to the matrix, other cells, etc. “It’s amazing what you can do with just the cells,” he said. His experience shows that “there is a synergy because you are creating environments.”

Stevens said that he was not sure that what Hartung said about mechanisms was achievable. He noted that when he entered graduate school, Gillette and Brodie had just published their covalent binding hypothesis in *Proceedings of the National Academy of Science*. The paper described the involvement of covalent binding to protein as a mechanism that caused both bromobenzene and acetaminophen hepatic toxicity in mice. They also demonstrated the protective role that glutathione played. Glutathione depletion and covalent binding – to steal a phrase from the environmental risk assessment world – are really key events in the mode of action of cell death for both of the compounds, Stevens said. “I think we could introduce this into the microphysiological systems today. Will the microphysiological system ever absolutely replicate the mechanism of pulmonary edema in a human lung?” he asked. “Probably not. You’ll probably find some margin of error in that,” he continued. “However, is the physiological mode of action in pulmonary edema replicated well enough in a microphysiological system that you can use it to make very, very good decisions about either therapy or safety?” Stevens argued that it could be very useful to introduce this idea someplace between the “mechanism versus correlation continuum”. “We tend to look at mechanism as a very absolute value – there is very little margin of error. But we still need to make decisions based on adequately characterizing systems that have a mode of action sufficient to advance compounds or remove compounds that are unsafe.”

In response, Hartung commented that “we live in a very pragmatic world. We will not be looking for perfect tools but for something that’s coming as close as possible to make reasonable decisions or at least avoid repeating the problems of the past.” Whatever system is devised will not be perfect. Hartung noted that his current perspective comes in part from an experience he had after leaving the European Center for Validation of Alternative Methods (ECVAM), where he served as director for six years, and before beginning his position with CAAT. The European Commission tapped him to spend a year focusing on homeland security issues. “What I learned was that a tremendous difference was made by moving away from all of the unknowns, which are possible threats, to think about what is actually damaged.” Focusing on the notion of critical infrastructure helped to focus their efforts. There are 50 ways to leave your lover, but there are only a couple of 100 ways to harm a cell, Hartung quipped. “If we follow this thinking that there are only a few Achilles’ heels in the cell, we will have a lot of angles covered because we can start looking differently into mechanisms. We don’t need to understand the entire mechanism, but we need to understand what is the way of bringing the system down,” he said. Hartung predicted that taking this approach would dramatically reduce the number the mechanisms that scientists need to contend with. It will also help us to focus on those which are really in the end harming us. “This is a different way of looking at the problem,” Hartung summarized.

Stevens countered that Hartung's argument was based on the mode of action "which gives you a correlation with an outcome, not necessarily a real mechanistic understanding," he said. "It's a sufficient mechanistic understanding, if you choose your key events which have a high correlation with either the presence or absence of an outcome you either seek or seek to avoid," he said.

Hartung responded: "Mode of action is a 3000-foot perspective on things instead of the molecular defined one," he said. "I think that we are underestimating here the opportunities given by modern technologies. Transcriptomics, metabolomics, proteomics and all of the other 'omics are giving us opportunities to molecularly phenotype things. Modern bioinformatics gives us tools to mine this and identify the signatures and common pathways," he said. The only thing that's stopping scientists from capitalizing on this more effectively is that no-one has yet figured out a way to put several gigabytes of data into something which in the end says "toxic/non-toxic" or "effective/not effective." "These tools are enabling us to do something different to mode of action. The focus is molecular phenomena," he said.

David Pamies of Johns Hopkins asked Kolaja about producing populations of different types of iPS-derived cells of sufficient purity.

Kolaja responded that, depending on the cell type, there are genetic selection markers that are put in so you have a cell type with low impurity, then you purify that cell type of interest. "You're still manufacturing a less-pure population of the cell. It's more about the aspect of identifying a selected purified population as the end result," he said. The interaction presents a bigger concern, though, speaking generically. It varies for every tissue type. "I think the key thing in the end is the microenvironment." Up until very recently work has largely been conducted on 2-D matrices. But as scientists move to 3-D cultures and better microenvironments, they will be able to address more complex questions about what types of cell populations need to be put back in to best replicate physiological effects, he said.

Pamies also asked about the iPS cells that Cellular Dynamics is making for clinical applications.

Kolaja said that they have a human pluripotent stem cell model that recapitulates the rapid differentiation toward a very primitive stage. It gives a very good ability to predict teratogenicity. What the scientists do with it depends upon what question they're trying to answer and their understanding of their disease model, he said.

Ingber commented that it is "really cool to get a few cells to build an organ. But you end up with something that's as hard to analyze as an organ. What you really want to be able to get mechanistic insight, is something that has the functionality and has the cell types but can be accessed in real-time at high resolution. That's what we're trying to do with our systems." Other people may be trying to build organs or organisms, which is something he has done in his own lab for 30 years, he said. But they are unlikely to be good for probing transport across cell layers, absorption, or the molecular basis of drug toxicity. What you look at depends on what question you're trying to ask, he said.

Stevens chimed in that biological network analysis has a great deal of value for prioritizing potential sources of inquiry using these methods. Weighted gene-code network analysis has great advantages because of the reduction in the dimensionality of microarray data while retaining all of the biology. We have 285 responsive networks in the liver, and 70% are either moderately or well-preserved in culture.

Ingber added that models have been improving quite a bit recently. The first gene-array models were called "guilt by association models" because all they showed were correlations, he said. But newer models are giving more mechanistic predictions, and they may prove critical for validating microphysiological systems. He said that he believes DARPA was very forward-looking because their applications for the microphysiological grants use computational approaches and informatics with human, animal and model data. His organization's grant includes a company called CFDRC, which he said does amazing multiscale modeling. The models can go from microarray to organ to whole body. They're developing models for the Wyss group's chips and seeing how they can scale back and link to the model, Ingber said.

Hartung noted that his PhD project with liver cells taught him the importance of co-culture, organotypic culture designed to replicate physiology. In his first project as a postdoc researcher, he got money from the European Commission to add perfusion to cultured liver cells. This was 20 years ago, and he was forced to find a way to supply nutrients and oxygen to the cultured cells. "It taught us the value of culturing things in an organotypic way based on how we expect the system to behave," he said. "If we use convenience as the only measure for evaluating our experiments – what we can do cheaply or quickly or easily replicate – it leads us to the shortcomings of current culture systems," Hartung continued. He suggested that we need to embrace these two components of co-culture and perfusion to make our models more organotypic.

Mike Holsapple of Batelle Memorial Institute echoed Stevens' suggestion of moving toward a mode of action as a way to begin to make sense of animal data vs. the human data vs. where the microphysiological organ systems are positioned. He said he believes that the decision framework treatment is very systematic. Holsapple asked Hartung if the mode of action framework precluded the analysis of molecular based mechanistic type data. "You talk about these pathways of toxicity… but I always thought that was another way of looking at a mode of action human relevance framework." Holsapple asked for guidance in understanding the difference between the mechanism and the mode of action.

The distinction between mechanism and mode of action arose mainly from the discussions around dioxin and other compounds evaluated through EPA's Integrated Risk Information System (IRIS) program, Hartung replied. "The difficulty of defining the molecular mechanism led to a compromise of defining a more cloudy mode of action," he said. The concept of the mechanism is considered to involve more precise knowledge of chemistry and molecular biology, while the mode of action is typically a more narrative type of framework. This implies that technologies that are already on the molecular level will be more likely to help us with mechanistic assessments," he said. His view is that the mode of action mainly helps where there is a lack of knowledge. "The terminologies are absolutely not clear," Hartung continued. "People are using them interchangeably." He said that it can make sense to interpret any type of molecular finding in the framework of mode of action in this setting.

5 Building representative microphysiological organ systems

Danilo Tagle, the associate director for special initiatives of the National Institutes of Health's National Center for Advancing Translational Science, chaired this session, which highlighted efforts to build a representative microphysiological organ system. He introduced the session, which focused on three organs important for understanding how the body deals with toxicity: the liver, the kidney, and the gastrointestinal system, including the microbiome.

Tagle also gave the audience some background on the federal government's efforts to develop organs-on-chips for drug screening via the Microphysiological Systems Program. The program's goal is to develop tissue-based chips that mimic human physiology to screen for safe, effective drugs using "the best ideas in engineering, biology and toxicology," he explained. In addition to the National Center for Advancing Translational Science, the program is being funded by the DARPA and the FDA. All of the speakers in the session are receiving funding through this program.

The chips that these researchers are constructing are ultimately intended to be combined. They will incorporate up to 10 organ systems using cells that remain viable for at least 4 weeks. The integrated systems will predict the human *in vivo* efficacy for new drugs, as well as their toxicity and pharmacokinetics. Their multicellular architecture will include vascularization with arterial and venous systems. They will also feature innervation, hormonal, humoral and immunological signaling. They will include genetic diversity and represent both normal and disease phenotypes.

Donald Ingber, director of the Wyss Institute for Biologically Inspired Engineering at Harvard University and principal investigator on one of the DARPA-funded projects, summarized why the federal program is important. The fact that the microchips reconstitute organ-level functions, as opposed to just cell or tissue functions, is key to their promise of accelerating drug development and replacing animal testing. "The problem, from our perspective, isn't only toxicity, it's the whole drug development pipeline," he said. Explaining why it is important to bring new technologies to bear on drug development, Ingber described Moore's Law, which holds that computer power doubles every 18 months and helps explain why microchips are now able to be used for such an application. "The price for microchips also drops every year," Ingber noted. In contrast, Eroome's Law – or Moore's Law spelled backwards – observes that the number of medicines invented halves every 9 years (Scannell et al., 2012). The price of drugs is also going up astronomically, he said. "The problem with current animal-based approaches to drug development isn't just the low percentage of drugs that succeed – there's a disconnect between results obtained in animals versus humans regarding molecular mechanisms and toxicity. And there's a great need for human disease models."

Testing a single drug compound can cost $2 million, Ingber pointed out. The studies take years to complete, innumerable animal lives are lost, and the results often do not predict clinical responses, he pointed out. "If I did this in my academic life – if I had a situation where I was wrong… even 30% of the time, people would think I'm nuts. But this is how the whole industry works. There has to be a change."

5.1 Liver microphysiological organ system

D. Lansing Taylor, the director of the University of Pittsburgh's Drug Discovery Institute, came to the podium to give the first talk on the microphysiology systems being developed with funding from the federal program. Taylor's university is collaborating with the Massachusetts General Hospital's Center for Engineering in Medicine to create a liver-on-a-chip.

Taylor reminded attendees that the National Research Council's 2007 *Toxicity Testing in the 21st Century* report articulated a major rationale for developing new systems for evaluating toxicity. At the time, most toxicology studies were not performed until a drug reached the pre-clinical phase of development. The report observed that this was very expensive and time consuming and exhibited a high failure rate, in part because animal models are not very predictive of many human toxic liabilities. It suggested that using *in vitro* tools earlier in the process might improve this situation.

Soon after, Taylor was involved with a company called Cellumen, now part of Cyprotec, which built a 2-dimensional, *in vitro* system called CellCiphr using hepatocyte liver cells. The scientists used the CellCiphr system to make safety alert predictions by building a database using dose/response information for both known toxins and unknown compounds to correlate a rank order for compounds likely to present safety problems. "At this point in time it was not powerful enough, although it was better than what people were doing… with the existing cell-based assays," Taylor said.

Taylor's team was inspired to try to improve upon CellCiphr's performance by creating 3-D systems that replicated the liver's acinus and sinusoids. These are the areas where nutrients, fats, toxins and bacteria that enter the liver via venous blood from the gut are processed. Once blood from the portal veins enters the liver acinus, it is channeled into the sinusoids where the blood is exposed to portions of the hepatocyte cells that secrete proteins and collect toxins for excretion into the liver's bile duct. Macrophages also bind and destroy bacteria and cell remnants.

Taylor's team's goal was to capture physiologies with the cellular mechanisms of action in these portions of the liver, including apoptosis, oxidative stress, mitochondrial dysfunction, phospholipidosis and cholestasis. They also wanted their model to be physiologically relevant and therefore include read-outs of important measures of liver function, such as the production of albumin and urea, leakage of lactate dehydrogenase (LDH), and secretions of glucose. To measure drug metabolism, they wanted to be able to simulate both phase I and II of the body's detoxification processes. Phase I consists of oxidation reduction and hydrolysis catalyzed by the cytochrome P450 enzymes that reside on the hepatocytes' membranes. In phase II, the fat-soluble output products from phase I are further processed so they can be eliminated from the body.

The researchers also set out to be able to measure and quantify the zonation, or levels, of oxygen in various portions of their biomimetic liver platform, as well as the pH of different parts of the liver. They also wanted to measure their liver's production of the bile that aids the small intestine in digesting fats.

The researchers aim for their device to reduce drug attrition rates by recapitulating the physiology of the liver acinus and using the results of testing for predictive database modeling.

The 3-D liver chip that Taylor's group has created has a grooved design intended to mimic the hepatic chords in the acinus. The channels created by the microgrooves represent sinusoids. Each sinusoid groove contains all the essential cell types found in the liver, including Kupffer cells, stellate cells, endothelial cells, and hepatocytes. The device includes sentinel cells, human cells that have fluorescence-based biosensors incorporated. The chip uses microfluidics to control the environment and the flow rates of medium continuously bathing the liver chip.

The group's goals for the 3-D liver platform include reducing drug attrition rates by recapitulating the human liver acinus physiology and making the optimal measurements to characterize it. A key element of this is building a predictive database, Taylor explained. "We also want to use all of the data we're collecting for a variety of systems biology modeling tasks," he said.

The microphysiology database is the heart of the system. It starts with documenting the protocols for assembling the liver device. This includes characterizing both the primary human liver cells and the sentinel cells used to report specific cellular activities placed in the device. Sentinel cells containing the biosensors make up about 10-20% of the cells in the device. It also includes measuring the device's characteristics, such as the flow rate. All of the information about both the device and any given experiment is captured. This includes the biochemical readouts and the mass spectrometry analysis.

Taylor and his group are also trying to take advantage of all of the databases in the public domain, as well as available drug data.

At this point, Taylor says that his team is isolating primary cells for use in the device from sources such as human liver resections, but they're also using human cell lines. They are also exploring how to use iPS cells. In all cases, they're both sourcing cells from outside sources and purifying them themselves. In his opinion, the "platforms will not be a valuable tool until we get to the point of having robust cells that are very reproducible." He stated his belief that iPS cells are most likely to help the scientists realize this goal.

The team began with plate studies to optimize the culture conditions and characterize the biosensor functions. Plate studies of up to 4 weeks showed that the cells are "happy" he said. Monitoring the secretions of albumin and urea showed that perfusing the cells with blood improved the production of both. They also characterized the chemicals produced during detoxification and metabolism.

The next stage is building sentinel reporter cells based on fluorescence-based physiological sensors. They are using lentiviral delivery with the option of homologous recombination to do this. They are also using simplified microfluidic devices to look at the effects of flow rates and how to optimize the microfluidics.

The group is testing and validating a number of biosensors that can be used in fluorescence-tagged primary hepatocyte cells used as sentinel cells. A simple example is fluorescent-protein-tagged cytochrome C. He showed a false color image showing that the mitochondria of primary human hepatocyte sentinel cells exposed to 10 µM of Nefazadone, an antidepressant that can cause severe liver toxicity, released the tagged cytochrome C from the mitochondria that is an early step in apoptosis. The group is using the Strictly Standardized Mean Difference metric to validate the biosensors, and any score above a 2 is considered acceptable.

Taylor also played movies showing how the biosensors work. One showed how the sentinel cells responded to exposure to 50 µM Menadione, a synthetic version of vitamin K3 that has been shown to cause oxidative stress in liver cells. The biosensors show that the exposure first releases cytochrome C and then activates the caspase 3 pathway, which is associated with apoptosis, or programmed cell death.

Other movies showed how the mitochondria in the hepatocytes took up calcium after being exposed to 50 µM Menadione, and how the same exposure resulted in a time-dependent increase in reactive oxygen species. A final movie showed how hepatocytes transported dye into bile canalicular spaces, the area where secreted bile is collected. The scientists used 50 µM Troglitazone to show that the transport could be inhibited.

As the scientists move to iPS cells, one of the things they are discussing is how to incorporate the biosensors into the cells by homologous recombination or via a lenticular delivery system, Taylor said. In the long run, Taylor's group hopes to make their biosensors commercially available, he said.

Taylor said that his group intentionally designed their microphysiology database so that it is organ, cell, and device agnostic. The design makes it possible to integrate external drug and target data and platform readouts for optimizing reference drug selection. It links to external databases to access information readily. It also has a means to store the complex sets of data that the platform generates, including cell and organ characterization for interpreting bioactivity results. All of this data on inputs and outflows can be used for systems biology modeling and creating predictive models.

The 120 liver reference drugs that the group used included 30 known to be hepatotoxic, 30 black-box labeled, 30 known to be toxic to other organs, and 30 nontoxic drugs. Over time, the scientists expect to increase the size of the database, Taylor said.

The microphysiology database is designed to store, serve, analyze and predict, Taylor said. The application is designed so that it can be easily installed and redistributable. The database will also store information about the layouts of the different devices, which may change. It is designed using a public API so that anyone can interact with it, and the design allows bidirectional interactions.

Both the database's backend and frontend tools are based on open-source software. It will be web-based. Taylor said that in order for it to be useful, he and his colleagues think that there will have to be strong administration, as well as different permission levels and user groups based on organs.

Taylor demonstrated how information about new compounds can be added to the database. He showed how users could search for the existence of test compounds and search publicly available databases to find data and then automatically populate the database. Another demonstration showed how new microdevice layouts can be designed. “This is like laying out a plate map for high throughput screening in microplates, to define domains within the device,” he said. It is important to show where measurements are being made from, he explained. The layouts can also be used to manage the uploading of the data from various measurements, he said.

After the team finalizes its 3d human liver sinusoid, the next challenge will be to integrate it with other organ devices, Taylor concluded.

5.2 Microphysiological systems being developed at Harvard’s Wyss Institute

Don Ingber gave his second talk of the day on the efforts of his group at Harvard’s Wyss Institute for Biologically Inspired Engineering to build representative microphysiological organ systems. He focused mainly on his group’s successes in constructing models of the human gastrointestinal system, including the microbiome, but he also discussed research underway toward recreating a human kidney proximal tubule, a small airway, and bone marrow-on-chips.

Ingber pointed out that his team also developed their lung-on-a-chip, which he described in the morning, as a toxicity model. After using the model to show how chemotherapy with IL-2 can cause pulmonary edema, they studied the model further to observe fibrin clot formation in real time. They also used the model to evaluate the artificial lung’s tissue barrier integrity after noticing that gaps developed between cells in the epithelium and endothelium. With the addition of physiological breathing motions and IL-2, they detected a four-fold increase in pulmonary vascular leakage, which drives pulmonary edema. “This was never seen before, or even thought about before,” he said. The Wyss team was able to confirm the finding of the new molecular mechanistic insight using an animal model.

Ingber told the audience that his group was able show a few years ago that the fastest molecular signaling that is induced in the endothelial cell is activation of the TRPV4 ion channel, which induces calcium flux in 5 milliseconds. He said that his group was aware that Glaxo Smith Kline had been working with TRP4 antagonists, and they were able to convince the company to collaborate to obtain some of the inhibitor. They were able to show that the inhibitor was able to completely prevent pulmonary edema in the lung-on-a-chip model. The model showed the same response as the company’s tests on dogs and rabbits with cardiogenic pulmonary edema (Huh et al., 2012).

The lung-on-a-chip model also revealed that immune cells are not needed to generate pulmonary edema caused by IL-2, and that breathing is responsible for a big part of pulmonary edema, which is relevant to patients on respirators. So the lung-on-a-chip functions as a disease model, drug efficacy model and toxicity model, Ingber summarized.

Ingber’s co-PI on his DARPA grant and Wyss faculty member, Kit Parker, has also developed a beating heart-on-a-chip. The device capitalizes on microcontact printing techniques to pattern extracellular matrix proteins to position and align cells. It uses primary cardiomyocyte heart muscle cells that are engineered to be cantilevered so that the cell contractile stress can be measured by bending of the substrate. The device also has been engineered to allow electric conduction to be studied, and the scientists have used it to induce and study arrhythmias. The results of testing contractile stress with this model were very similar to those produced by the rat ventricular stress tests regularly used by pharmaceutical companies.

The Wyss researchers received one of the first grants supported by FDA’s Advancing Regulatory Science program to link the lung-on-a-chip with the heart-on-a-chip. Since then, they have developed a way to deliver aerosols directly to the chip that produces 1-3-micron-sized aerosol droplets.

The Wyss researchers have recently begun to outsource the production of their organs-on-a-chip, and are moving toward mass production. They have recently signed an agreement with Sony DADC, and the company predicts that disposable organs-on-a-chip could be produced for a few dollars per chip in the future if the market is large enough. These predictions are made based on the lung organ chip, but Ingber says that there are many similarities between the organ-on-a-chip models that his Institute is developing. This is in part because almost every organ system involves a tissue-tissue interface between vascular cells and parenchymal cells that provide specialized tissue functions. Flow and mechanical deformation are also crucial to being able to study the organs on the chip.

The Wyss researchers also have produced what they call a peristaltic human gut-on-a-chip, Ingber said. They started with the lung organ and modified it to mimic the human intestine with its microbiome. They made the organ higher and wider and set it up to produce a trickling type of a flow similar to that in the human gut. Rather than exerting breathing motions, it was engineered to have the cyclic deformations associated with the wave-like contractions of peristalsis.

They used human CaCo-2 colon cancer cells, which are well known to not be well-differentiated, and in existing static culture systems they appear more like skin cells than gut cells. However, after just three days of experiencing the flow and strain in the artificial gut-on-a-chip, they began to look columnar, like gut cells. Over time, they spontaneously reorganized to form villi, which are the cell-lined finger-like projections that are normally found lining the human gut. Just like the villi in the human intestine, these structures have tight junctions and are covered with mucus. The structures also include the crypts containing proliferative cells found in the human gut, which include four different types of differentiated epithelial cells (absorptive, mucus-secretory (Goblet), enteroendocrine and Paneth) that take characteristic positions similar to those observed in the living human small intestine.

Ingber’s group has also used primary human intestinal epithelial cells to recreate the same structure with the gut on the chip, as recently described in articles in *Lab on a Chip* (Kim et al., 2012) and *Integrative Biology* (Kim and Ingber, 2013). The fact that colon tumor cells work so well to reproduce the intestinal physiology and even produce mucous suggests that the cells are “rebootable,” especially since CaCo-2 cells are known normally not to produce mucous in static cultures. Transcriptome

profiling of the gut epithelium revealed that the expression of about 10% of 22,203 human genes was significantly altered in mechanically active environments including flow and/or strain. Just trickling flow can alter the phenotype of these cells, and trickling flow plus cyclic strain change them in an entirely different way, Ingber said. To him, this suggests that "there are no bad cells, just like there are no bad kids – they just get in a bad environment."

Ingber said he began the project to reproduce the microphysiology of the human gut about five years ago when he learned of the key role played by the microbiome. Thus, the Wyss team began adding living microbes to their gut-on-a-chip. They began with Lactobacillus GG isolated from humans, which increased the gut on chip's barrier function. In contrast, if the Lactobacillus bacteria are added to the same gut cells cultured in a static Transwell culture plate, the cells die in 24 hours.

The Harvard Wyss researchers' also have explored development of a human kidney-on-a-chip, as described in their recent *Integrative Biology* article (Jang et al., 2013). This effort began with a focus on the kidney's proximal tubule because of the key role it plays in drug toxicity. They started with a simple, inflexible system to evaluate the effects of the shear stress on the tubule physiology. They used primary human kidney proximal tubule cells, and found that with tubular flow, the cells became more columnar. The epithelial cells also become more polarized and increased formation of primary cilia, the slender protuberances on the surface of the cells involved in chemical sensation, signal transduction, and control of cell growth. The cells' re-uptake of albumin also increased, which is important for normal kidney functioning.

To evaluate how the proximal tubule-on-a-chip functioned in response to a toxic substance, the Wyss researchers used cisplatin, a chemotherapeutic agent known to be toxic to the kidney. Their results showed that the model with tubular flow did a better job of demonstrating how cisplatin affected both the release of lactate dehydrogenase (LDH), which indicates cell injury, and apoptosis than static cells. The chip-based model also responded when a cisplatin inhibitor was used, unlike the static cells.

The researchers were able to reconstitute the function of P-glycoprotein (Pgp) transporters, which are found in proximal tubular cells and play an important role in both drug toxicity and drug resistance. "It's the first time in a human kidney microfluidic system that we've been able to see enhanced functionality of a toxicity response," Ingber said.

More recently, the team has been working on a model that goes beyond the alveolus-on-a-chip to mimic the diseased airway associated with chronic obstructive pulmonary disease (COPD). They have been able to demonstrate differentiation of primary human bronchiolar epithelial cells on their chip. The chips include cilia that actively beat in the microchannel. Further, they have been able to simulate bronchiolar inflammation with the chip device. The researchers activated the small airway-on-a-chip system's toll-like receptor 3 (TLR3), which plays a fundamental role in pathogen recognition and activation of innate immunity.

From there, the researchers mimicked a viral infection with their small airway chip using poly I:C (polyinosinic-polycytidylic acid), a synthetic analog of double-stranded RNA (dsRNA), and demonstrated increases in multiple cytokines that are representative of what has been seen in humans and animals, Ingber said.

The Wyss team had only been working on this project for less than 9 months in May 2013. For Ingber, this demonstrates that "these things can work quite quickly, more quickly than people might expect."

Wyss researchers are also developing a microphysiological model replicating bone marrow-on-a-chip. The model has potential for bone marrow transplantation, *in vitro* blood cell manufacturing, and as a stem cell niche model.

The approach they used is "very different and much more complex," Ingber said. For stem cells, particularly bone marrow stem cells, the concept of the niche is very important, he said. The researchers began by putting demineralized bone powder and bone morphogenetic protein into a microdevice that they implanted under the skin of a rodent to induce bone to form. After 8 weeks they removed the device and the bone and put it in a microfluidic device, punched holes in it, and started perfusing it.

They found that the distribution of hematopoietic stem cells in the synthetic marrow was identical to the distribution in mouse bone marrow. The red and white blood cell distributions were also identical, Ingber said. The Wyss team is currently in the midst of translating the concept for human cells, and it is very promising, Ingber said.

Ingber also has a vision for integrating the organs-on-chips to produce what might be described as a human body-on-a-chip. He proposes to do this by using microfluidics to connect the vascular channels of the different organ chip devices. This will produce an organ circuit that can, for example, show how an aerosol-based drug would be delivered to the lung, absorbed into the circulation, metabolized in the liver, excreted by the kidney, and determine whether or not it will produce toxicity in the heart. There is also the potential for having blood plasma flow through the organs on the chip. "We're really moving towards having an automated instrument that pharmaceutical companies can use," he said.

The Wyss Institute has received funding from DARPA to produce an instrument for integrating and interrogating the organs-on-a-chip. It is intended to be "plug and play" like a DVD player, Ingber said. The institute has designed a universal cartridge holder that can be plugged into an organ chip and used to insert it into the analytic instrument, as well as to download data on biochemical parameters such as levels of oxygen, glucose, lactate and pH. The instrument will have integrated imaging capabilities and will permit sampling for mass spectrometry analysis as well. Wyss is developing software to control the instrument as well as use the data collected to predict human pharmacokinetic and pharmacodynamic (PK/PD) responses.

There is still more to do, but "if we keep the progress up, we really think that we might have the beginnings of an instrument that pharmaceutical companies, cosmetics companies, and regulators can begin to use to evaluate if it could be helpful," he said.

Ingber's questions from the audience included one from Myrtle Davis, the Branch Chief for Toxicology and Pharmacology in the National Cancer Institute's Division of Cancer Diagnostics and Treatment Developmental Therapeutics Program. She asked for interpretation of Ingber's group's findings about IL-2, a cytokine which is prescribed to some cancer patients to activate their immune systems to fight the cancer. She said that it's exciting to show that IL-2 will bind to its

respective receptors and cause capillary leaks in culture, which is one of the syndromes that occurs in patients being treated with the drug. However, the number of patients who experience this side effect is somewhat low compared to the number of patients with renal cancer and melanoma who have experienced positive side effects, she continued.

Ingber said that he could not yet determine what percentage of patients would be likely to benefit or have the side effect of the treatment or if the dose/response is different under different conditions.

Davis responded by asking if having the other cells there might have been the mitigating factor with respect with whether the toxicity becomes an active, irreversible toxicity that is not manageable. Ingber said that this is not yet clear, but "the point is, from our side, that we can ask those questions, such as what happens with or without immune cells present" and see how the situation changes under different conditions. "We need to criss-cross the results we get using informatics and modeling with what we know about humans and what we know about animals," to make detailed evaluations, he said.

Mel Anderson of the Hamner Institutes for Health Sciences asked about the design criteria for different organ systems. Ingber said that his collaborators at CFDRC Corporation are developing PK/PD models of whole living organs and scaling them to try to get an "elemental scaling" for each organ's functional unit. The architecture is a big challenge, but it presents a problem we think we can address because we can quickly manipulate system design and carry out iterative testing, he said. He credited DARPA with doing a very good job of handling the requirements.

Anderson also commented on tissue flow in the organs-on-a-chip. He pointed out that the plasma volume and the liver volume are almost identical and asked about the challenge of designing that into the system. Ingber responded by agreeing that it represents a big challenge, but he stressed that the goal is not to rebuild an organ or organism. Instead, the goal is to build models of minimal organ functional units with known limits, he said. "We can't do everything with these simple models, but we can do a lot better than we can with cells, and most importantly, we do it with human cells" he concluded.

5.3 A second kidney proximal tubule-on-a-chip

The final talk of the session was by Jonathan Himmelfarb, MD, the director of the University of Washington's Kidney Research Institute, and a professor of medicine there. He talked about his group's efforts to create a kidney-on-a-chip.

Over the past 25 years, the incidence of end-stage kidney disease and acute kidney injury has been rising steadily. Over 20 million Americans now have kidney disease, and drug toxicity is a major source of the disease, Himmelfarb pointed out. Often after acute kidney injury there is incomplete resolution, which contributes to the epidemic of end-stage kidney disease. "It's really only in the last couple years that we've realized that toxic kidney injury contributes substantially to end-stage kidney disease," he said.

The rate of blood flow through the kidney is approximately 20-25% of cardiac output, Himmelfarb said. The kidneys play a primary role in the elimination of 20-25% of drugs and their metabolites. The organs are highly susceptible to injury from drugs, particularly the proximal tubule. People with kidney disease have altered metabolism and elimination of drugs, and their risk of having an adverse drug reaction is greatly increased. Up to 20% of hospital admissions for community acquired acute kidney injury are attributable to drug induced kidney injury.

Himmelfarb said he is hoping that the microphysiological platforms will lead to new therapeutics, and that they will have a role not just in a better understanding of toxicity but also efficacy. He displayed a list showing 8 new drugs for treating kidney disease approved by the FDA over the last decade. Many are "me, too," drugs he said, and not a single one is able to substantively change the outcome for the majority of people with kidney disease.

The formula for how the kidneys clear drugs and toxins was first laid out almost 100 years ago. The rate of filtration and the ability to reabsorb through the kidney's nephron and tubules is well-understood and both can be easily modeled clinically. "But to this day we cannot effectively model tubular secretion of drugs and toxins," Himmelfarb said. To the pharmaceutical companies, tubular secretion remains a black box in terms of understanding either if any compound is secreted or the extent to which it is going to be secreted, he explained.

Himmelfarb showed an image of a normal human kidney indicating the tubules, the vessels, and the perivascular cells. The vessels wrap around every tubule. In the last couple of years scientists have learned that not only are the perivascular cells important for function, in terms of secretion, but research has shown fairly conclusively that perivascular cells are the disease-causing cells that contribute to fibrosis in kidney disease.

A paper published online the week before the conference shows seven-fold differences in the rate of clearance for compounds secreted via the same organic anion and cation transporters within the kidney, Himmelfarb said. In dialysis patients the rate of clearance ranges from 3-fold to up to 100-fold of the levels of healthy individuals. "To this day, nobody can really model for a given compound how it will be cleared by people with less severe levels of kidney disease," he observed.

The functional unit that Himmelfarb's group set out to model includes the vasculature of the peritubular capillaries with the pericytes, which communicate with the blood vessels' endothelial cells, and the proximal tubule. This unit is critical to kidney toxicity and how the kidney eliminates drugs. The proximal convoluted tubule cells are full of mitochondria and are highly active metabolically. "The work of the proximal tubule is all about transport, whether it's reabsorption or secretion. It's really the factory in the kidney for the transport of solutes," he said. Its functions also include the generation of ammonia from glutamine, or ammoniagenesis, and the 1-α hydroxylation of vitamin D. It is also the cell in the kidney that is most subject to injury because it is so metabolically active and is exposed to such high concentrations of the kidney filtrate, Himmelfarb said.

By including three types of cells, the proximal tubules, the pericytes and the microvascular epithelium, Himmelfarb said his group's tubular interstitium-on-a-chip should be able to effectively model the kidney's secretory process. The chip they are designing includes a parallel tubule and a parallel microvessel.

The platform that the group is working with primarily is made by Nortis, a start-up company that emerged from the University of Washington. It makes small, disposable, relatively low cost chip-like devices for the creation of vascularized 3-D microenvironments of human tissues and organs. The chips are traversed by one or more tubular cell structures. The technology allows for compartmentalized flow through both lumens, and the extracellular matrix from outside the cells to perfuse around the proximal tubule interstitium independently.

There are no artificial surfaces to which the cells must attach, and the design allows for the scientists to control the shear force of the luminal liquid flow, Himmelfarb said. The device is also microscope friendly and easy to clean via an autoclave.

Himmelfarb's group began by using primary cells obtained from individuals who were having all or part of a kidney removed via nephrectomy for a small, circumscribed cancer where the rest of the kidney is normal. The 3-D microphysiological system they have constructed with these cells can now reproducibly function for 28 days, but Himmelfarb acknowledged that his group is continuing to work on quantifying the chip's performance.

In the next six months, the team plans to characterize the chip's glucose reabsorption, its secretion of organic anions, such as para-amino hippurate, and its expression of kidney injury molecule-1 (KIM-1). They also plan to characterize the model's synthesis of glutathione, its production of ammonia, and its secretion of organic cations such as creatinine.

On the microvascular side, Himmelfarb's group has been isolating human kidney microvascular epithelial cells. "To our knowledge, we are the first group who has ever been able to do this successfully," he said. They are able to show with CD31 expression that the cells develop junctions in primary culture. The group has used PCR to show that the cells look like microvascular endothelial cells.

Himmelfarb said that he believes his group is also the first to isolate kidney-derived pericytes from human kidney so that they have an appropriate phenotype outside of the body. He showed a slide demonstrating that the pericytes express appropriate cell markers.

Himmelfarb's group has also cultured the pericytes with the microvascular epithelial cells and renal tubular epithelial cells in a microperfusion device and observed that they self-aggregate into microvessels. The junctions are tight, he commented, but it looks like there are holes in the epithelia and the vessels are leaky. "This is not unexpected," Himmelfarb said. With mouse kidney microvascular cells they have confirmed a previous report that the endothelium is indeed fenestrated with holes, and they intend to confirm the finding with human cells.

Pericytes are crucial for creating the microphysiology of the proximal tubule interstitium, Himmelfarb said. Co-culturing the pericytes with the endothelial cells in the microperfusion devices results in a tremendous increase in the production of extracellular matrix, he said.

Himmelfarb said that his group has spent a lot of time discussing what is the best initial source of matrix for their devices. They initially used a hydrogel coated with collagen, but are moving toward taking human kidneys and decellularizing them to produce a source of matrix for their system.

Many challenges remain to optimizing the group's kidney microphysiological system, he said. The group is still working out what the appropriate cell sources are, including exploring how well iPS cells might function. They are still experimenting with different approaches to cell seeding and adhesion, as well as the benefits of using organ-specific extracellular matrix. Other subjects of continued research include the best approaches to the kidney-derived microvessels, because it is not easy to get the cells into culture. They are also working on the flow dynamics and the analytical chemistry, as well as how to move from descriptive to quantitative assessment of the device's functionality.

The path forward towards FDA validation will be dynamic, Himmelfarb said. "We are taking a bottom-up approach," he explained. They are currently focusing on building the best possible human kidney tubule interstitium and optimizing it. From there, they will figure out how to merge it with other organ systems. They will clearly need consistent cell sources going forward, as well as consistent, high-quality real-time biomarker and imaging read-outs that are not destructive.

"I believe that these devices can and will have a role in enhancing our ability to predict kidney metabolism and elimination of drugs," Himmelfarb said. He said that his group will be using standardized test compounds in the next year to start to assess the model's functionality and ability to predict toxicity. He predicts that these devices will have a role eventually in personalized medicine in terms of genetic analysis of donor transporters and drug metabolizing enzymes.

In the short term, Himmelfarb said he thinks the real value of these devices will be their ability to produce high quality and high content data at key stages of the drug discovery process.

Over the past decade, he noted, there has been a significant effort of scientists to qualify biomarkers for predictive safety testing. The Predictive Safety Testing Consortium includes scientists from pharmaceutical, biotech, academic and regulatory organizations, including the FDA. The consortium was able to develop biomarkers, including KIM1, which consistently outperformed other measures of kidney function, particularly for detecting early kidney disease. Moving from identifying the biology to developing the assays to FDA qualification only took a relatively short period of time, 3-4 years, Himmelfarb pointed out. "They're now commercially available on a variety of different platforms." Similarly, he expects that the microperfusion systems hopefully will be able to move forward after further development.

Once it is developed, the kidney-on-a-chip may help scientists in a number of ways, Himmelfarb said. It may aid in understanding uremia, or the loss of kidney function. It may prove to be a tool for improving kidney organ preservation for transplantation. It may also help pharmaceutical companies develop new drugs to combat kidney disease. The tool could help scientists understand kidney development and even grow new kidneys via organogenesis. Eventually, it may even allow the creation of implantable artificial kidneys, he concluded.

6 Discussion about representative microphysiological systems

After the session's talks were over, Tagle moderated a panel discussion featuring Taylor, Ingber and Himmelfarb that included audience questions and comments. The first subject of discussion was personalized medicine.

Himmelfarb told the audience about a modeling application of personalized medicine which could show how different medicines can affect toxicity using cytochrome P450 3A5 as an example. It is enzymatically almost indistinguishable from CYP 3A4, and it is highly polymorphic. Only about 20% of individuals actually express 3A5 in the kidney, he said. Eighty percent of people have no CYP3A5 function. His group has not done genotyping, but has seen by staining that some kidneys have CYP3A liganding. The function of some of the potent immunosuppressing drugs in clinical use, like cyclosporine, is limited by nephrotoxicity, he noted. The plasma levels of cyclosporine do not always correlate with nephrotoxicity. He said that clinicians hypothesize that those who express CYP3A5 in the kidney probably do not experience nephrotoxicity because they will produce toxic metabolites intermittently that will not be measured in blood levels systemically.

Himmelfarb said that donated kidneys which express CYP3A5 could be studied via a pharmacogenomics chip which is focused on transporters, metabolites, and enzymes. His group's system enables them to see how perfusing different levels and amounts of cyclosporine will affect toxicity.

He reminded the audience that his group does not buy commercial primary cells. They use nephrectomy specimens obtained clinically at the University of Washington. He said that they see different growth characteristics based on donor age. "There are probably very important factors in terms of drug efficacy and drug safety related to things like telomere changes and senescence that they can examine with microphysiological systems, he said.

Ingber said that George Church of Harvard University Medical School's Personalized Genome Project has collected 16,000 skin biopsies. The donors sign a waiver allowing him to do full-genome analyses. "We think of personalized medicine as encompassing more than individual patients," he said. It can also help pharmaceutical companies change their paradigm. If they have a drug that fails in clinical trials, they may be able to find a subpopulation that responds and redesign the drug for that subgroup. "Or what if you could design a drug for subgroups from the beginning?" he asked. It might prove to be a more cost-effective way to do personalized medicine, he said.

Ingber's group is beginning to get primary cells from patients with diseases including asthma, COPD, and chronic inflammatory bowel disease. This gives rise to the possibility of eventually comparing cells harvested before donors have a disease and afterwards, he said.

Taylor said that the philosophy of the University of Pittsburgh's Drug Discovery Institute is not just to be a "mini version of the pharmaceutical industry but to implement some highly innovative things that might alter the way things are done in the future." He and his colleagues "can afford the time" because they don't have the pressures of trying to push a single molecule into the clinic. They are trying to integrate the "best of computational biology and chemistry and systems biology with more advanced experimentation." As a result, they are moving more and more towards 3-D systems at the early discovery stages to tackle problems like metastatic cancer. This requires engineering metastatic models. "It's a huge challenge, but we think a combination of tissue engineering and novel approaches to developing therapeutics will be worthwhile and a key to personalized medicine," he said. That is another place where iPS cells may have an impact, he added.

James Stevens of Lilly thanked the speakers for bringing up quantitative systems pharmacology. His question was about scalability and classification versus network. "I'm beginning to be a little skeptical about the ability of systems trained on compound classes as having any scalability across either vertically – or laterally – integrated complexity." He said that his impression, based on his company's preliminary investigations, was that researchers might be better off finding the biological response networks and then using the robust, scalable networks to classify compounds. He contended that it might work better than the practice of beginning with a set of compounds because it "introduces a bit of an artificial bias in how we classify them." Then they would try to train the system to recognize different classes. This approach also requires that a new system be rebuilt and retrained when you move to a different organ, he said.

Himmelfarb responded that over the past 25 years, the great pressure to get compounds out has minimized creativity and innovation. "I think the approaches you described are going to be critically important," he said. But he also said that he was not convinced that it would answer the question. "I think other paths are going to be important." He said that the approach that his group has chosen to take for efficacy is one where they look at more complex model systems with model compounds. They go on to develop the compounds that do better, he said.

Stevens responded by reframing the issue to focus on gene signatures versus biological response networks described by gene expression. "Our experience with gene signatures is that they are absolutely not scalable," he said. "You have to rebuild the signature in primary hepatocytes, you have to rebuild the signature in liver, and you have to rebuild the signature in heart." He said his group's experience is that biological response networks, if adequately described statistically, are much more scalable. "Now the information becomes more translatable, but you can still use them to do compound comparisons." He said that one of the challenges that the area needs to address is the scalability issue. The idea is that a system should not become unique unto itself in its ability to provide predictability. Instead, they would like to have a system that is stable across systems, both organ-to-organ and horizontally integrated across complex systems, as well as vertically from simpler cell-based systems to more complex innovative physiological systems."

Himmelfarb said that he didn't disagree at all with Steven's observations. "The challenge will be: What kind of read-outs are you going to base those pathway analyses on? Do you have thoughts on that?" he asked.

"If you collect any sort of biological read-out, whether it's the proteomic changes, the metabolomic changes, or the gene expression changes, each in its own way is a reflection of a biological response network. I like the idea of using statistics

within the system where you induce sufficient variability to look at the co-regulated sets of those individual elements. It doesn't require that we apply our bias. You can say, 'if the biology is true, the statistics should prove that, and then you can prove or disprove [the hypothesis]. I think you could start with any or all high-content data, but if you start with a robust statistical approach that depends upon the conservation of biological response networks, that will become stable. And then we build in the classification compound, beginning with what may be an overly lofty view – these all cause cholestasis, these all cause necrosis, etc. – when in fact they may cause a similar pathology, but the differences between those compounds may be more different than the similarities."

Himmelfarb replied that he thought this should be part of a major shift in the way drug discovery and development is performed. He asked Stevens how broadly it could be applied.

Stevens responded that they were just beginning to implement the system so they will know in the future how broadly it can be applied.

Next, Lewis Kinter of AstraZeneca asked what he termed a very basic question. "The three of you presented results in your version of the lab-on-a-chip, and you had a different cellular interface. Dr Ingber was growing cells on PDMS, Dr Taylor grew cells on some sort of matrix, and Dr Himmelfarb is growing them on collagen and maybe some sort of matrix. Can you discuss the differences and the potential advantages or disadvantages of growing cells on different matrices?"

Ingber responded that his team covers the PDMS with extracellular matrix and they grow cells on the matrix. "Everybody else has some kind of material. But many have pointed out that PDMS is a problem because a lot of small hydrophobic molecules will get absorbed. We just submitted a paper where we describe other potential materials. We have a few different materials now; they're clear, flexible, and don't absorb small molecules. As we work with Sony, before we scale up, we'll be making a final decision," he said.

He continued by pointing out that he has been a matrix biologist for more than 30 years. "What you want is a matrix that will get the cells in the right place and convey the mechanical signals and put them in juxtaposition. But they should do it on their own – in my experience, if you produce the right environment, they'll produce the right matrix on their own," Ingber said.

Himmelfarb labeled the matrix as critical. "Without doubt, over time, all of the cell types produce their own matrix," he said. But it isn't yet clear whether or not the initial conditions have a lot to do with self-assembly, differentiation, initial phenotype, or cell-to-cell communications that affect the way these devices function.

There are a lot of complexities, Himmelfarb continued. "For a lot of these devices, our strategy is, you want the cells to be dedifferentiated when we load them into these devices because they'll proliferate and make whatever shape we want. But then we want them to redifferentiate, form tight junctions, form functionality inside the devices over time under the right conditions – the right flow conditions, the right cell-cell interactions, the right matrix. There's an enormous amount of trial and error in this process – initial adhesion, what shear stress can you put in, what are the flow characteristics at the beginning vs. as that differentiation process goes on. Those are the kinds of things we've been learning by trial and error," he said.

Taylor said that he agreed with what had already been stated. "In particular, if you're going to have a system that goes out for 3-4 weeks, whatever initial matrix you put down is set to be gone. Creating the microenvironment, with the right cells and the right 3-D and the right flow to induce these cells to be the most active and create their own matrix. That's key, I think, to their physiological function."

Ingber added that the DARPA grant requires the microphysiological systems to be kept alive for a month. "NASA keeps things alive in space with perfusion cultures. You should be able to keep them alive as long as you perfuse them. I don't think long-term is a problem for us."

Another questioner asked the presenters when they expected to create an actual organ-on-a-chip, rather than subsections of tissues with different positive responses.

"They're more than subsections of tissues," Ingber replied. "They're organ parts-on-a-chip… I don't want to build whole organs. I had a company in the 1980s that built whole tissues. I thought I'd die before I'd see that actually had an impact. Whereas, we can get the minimal things we care about in these organ parts or functional units on chips. We need to figure out what we need for what we're asking."

Himmelfarb added his view that there will be insights from the microperfusion systems that will be relevant to efforts in organogenesis or regenerative medicine toward trying to build a whole organ. "Those are completely separate questions for completely separate purposes," he said. The systems will be most useful if they are focused on specific unmet needs in drug development and drug evaluation, because that is the goal of the NIH-DARPA program. For example, the design of the Himmelfarb group's kidney-on-a-chip was informed by the reality that it's very easy to model kidney filtration, but it's very hard to model tubular secretion. "We chose to focus on what we call a kidney tubule interstitium, although it is really a proximal tubular interstitium, because that is the critical unmet need for drug discovery and drug evaluation."

Ingber summarized, "You want a disease model that people actually care about. We're not trying to just do drug safety. We're trying to produce models that have different uses."

Himmelfarb continued that it is very difficult to get a holistic model through a drug safety evaluation because they are very finite evaluations.

Ingber added his belief that the models will be good for drug candidate lead prioritization and having insight into the molecular mechanisms of both the drug's action and any associated toxicity. It may not be used for baseline screening, but it may help identify the pathways that are important. "If there is a problem, it may enable researchers to make selections based on knowledge." Over time, that may feed into opportunities to use the systems in a more widespread fashion. "But we're not trying to build a whole human all at once, or trying to solve all problems at once." If we could get results better than what tests using cells from the FT2 human kidney clear cell carcinoma line on a dish can provide, that's an advance, he said.

Taylor commented that he was impressed with the collaboration between DARPA, NIH, FDA and EPA "to try to do something that on the face of it is almost impossible." This DARPAesque approach requires a long-shot view. He said that he has been impressed by what he had learned about how much has already been accomplished by trying to do the near-impossible from listening to his colleague's presentations over the last six months. "If we can get to the point, as Don [Ingber] said, of having some simple tools that can work better than the tools we have now. I think we're essentially there now." He predicts that how much the project will ultimately achieve will be dependent upon tissue engineering and systems biology. "But there's a lot that can be accomplished without us constructing a whole organ," he summarized.

Kinter said that he appreciated that the models use human tissue and are aimed to improve safety predictions. He noted that the pathway for getting a new drug into a clinical trial requires testing with rodent and non-rodent species. He asked the presenters how any of their systems will fit in with the 14- or 28-day rat and dog studies. What if testing with the microphysiological systems did not show a signal and the rat or dog testing did?

Ingber responded: "Life is short, and I think the goal is to get to humans." FDA's requirements for new systems are that they need to be as good as or better than what they are replacing. If the new microphysiological systems are as good as the animal studies, which will have to be done in parallel, and we get a different result, "we'll have to use informatics, and 'experience.'" Because he doesn't think a pharmaceutical company can use that information, "it may take a long time to figure out what was right or wrong, or we'll have to design new types of systems. We certainly can use rat and mouse cells in our models, but it's really a lot of work. It's much harder than it sounds when I presented it – it's incredibly complex."

"To design a system for what we're trying to get rid of does not make sense and I don't want to take my people to do it unless it's absolutely critical," he continued. Perhaps in a case like the one Kinter hypothesized this might make sense, he said.

Kinter said that if the scientists really are thinking about moving into the regulatory space and using these systems as replacement for animal studies, it will require building rodent and non-rodent homologues.

"It would be very easy to use rat or mouse tissues in the systems that have been created," Ingber said. "In the end, you're probably right that we're going to need to do that."

Luciana Borio, the assistant administrator of the Office of Counterterrorism and Emerging Threats (OCET) in the FDA's Office of the Chief Scientist, broke in. "We're trying to move the regulatory space to be closer to the available science," she said. "We forget that the reason why things are required today is not because they're written in stone and will be there forever. We're trying to evolve the paradigm based on new science."

Kinter said that he appreciated this. Then he asked if Borio could say how many regulations for preclinical testing have actually been removed over the last three years.

"I can't respond with statistics," Borio replied. "But I can tell you that this effort is a beginning to attempt to do that. There have been major regulations that have been modernized based on new science."

"The best example we've got of that is where we've shown that both industry and regulators no longer use the data from a particular test, such as a particular gene toxicity test," Kinter said. "When you can see that no-one uses the data anymore to make a decision, then it becomes a relatively simple process of withdrawing a regulation." Based on his years of experience with bioassay type endpoints, which are not necessarily mechanism-based and make correlations that are less than perfect, he said, the argument will always be: "Well it has good correlations for the compounds in your test set. How's it going to hold up for compounds you're making?"

Ingber responded that the new microphysiological tests will allow researchers to identify molecular mechanisms of toxicity. "It can be vetted to forward-looking pharmaceutical companies' research efforts to begin to identify… the molecular basis for their toxicology testing," he said. Additionally, he pointed out that Throckmorton discussed FDA programs that hold out the possibility to show that these chips are as good as an animal model. This raises the possibility that one by one we may be able to replace animal models with these systems for particular disease models and maybe particular toxicity models. "It's not going to be overnight, so there's no way we're going to replace whole-animal testing," he said. "But I think pharmaceutical companies need to realize that what's here today is not necessarily going to be there 20 years from now."

He added that he is encouraged because everyone is working together and educating each other. "Every time I come down here, I learn something totally new," he said. "Maybe the toxicologists will never accept this and it's not where we should be focusing. Instead maybe it should be disease models and safety testing for food or other things that have a lower threshold."

"There's always a challenge when there's a reference standard," Himmelfarb said. "I think we could agree that the animal testing for preclinical evaluations is the reference standard, but it's suboptimal for all of the reasons we've been hearing about." As an example of the difficulties associated with changing a reference standard, he brought up kidney injury biomarkers. For the longest time as they were being developed, they were being compared to the gold standard of a rise in sero-creatinine, which scientists considered a suboptimal reference standard. The lengthy process of replacing the standard involved an FDA qualification process to prove that the biomarkers were more sensitive than the reference standard.

"It is very early in the development of these microphysiological systems," Himmelfarb continued. He expressed confidence that they would all be tested over time against test compounds and existing reference standards. "If these microperfusion systems outperform animal models in predicting clinical toxicity, one would hope that they can become a reference standard and replace animal model testing in the future. That's not going to be an immediate process, and there's a lot of development that needs to go into these systems before we can get there. But I don't think there's any value to putting animal cells into these bioperfusion systems. You're probably going to get the worst of both worlds if you do that. I think that we should stick to human cells with the hope that we'll learn more about human toxicity plus genetic variation and personalized approaches to subpopulations that might benefit. You're not going to get any of that from inbred animal model studies."

Ingber told the audience that “the more drugs we can get that actually failed in humans but looked good in animals,” will be helpful. “If we can mimic that, it will be much more important and exciting than if we get a different result and don’t know what to do.” It would be extremely helpful to be able to show that the microphysiological systems can identify what the animal testing missed.

Kinter said that he felt that Ingber put his finger on a key issue. “What’s key is to figure out what we should have to show to come to the conclusion that you have outperformed the existing paradigm.”

Ingber said that his group is partnering with the pharmaceutical companies because these companies “have a better feel for what would impress you than we do.” The companies need to help us identify the appropriate benchmarks for doing better in terms of identifying both efficacy and toxicity.

Tagle concluded by pointing out that exchanges like the ones taking place at this conference are exactly why federal agencies like to convene workshops to bring together the stakeholders, including FDA, pharmaceutical companies and academic researchers to identify how to best work together to achieve common goals.

7 Integration of microphysiological systems

The final session of the conference focused on the integration of microphysiological organ systems. The session’s speakers presented their observations on what was essential, what was nice, and what wasn’t.

First up to the podium was Robert Kavlock, the Deputy Assistant Administrator for Science within the Environmental Protection Agency’s (EPA) Office of Research and Development. He began by saying that he was “absolutely amazed by the progress” in creating the microphysiological organ systems that the previous speakers described. He said that both he and his fellow speaker, Melvin Andersen of the Hamner Institute, are “interested in firming the foundation for risk assessment.”

Both scientists in the pharmaceutical research area and in the environmental health research area are motivated by the new technologies’ ability to “allow us to get inside the black box” between the exposure to a chemical and the responses that can be detected by animal testing, Kavlock said. But the scientists from the two disciplines are driven by different goals.

For EPA, the driver is not successful clinical trials. Instead, the agency is driven by the need to collect toxicity data, Kavlock said. There are too many chemicals and too little data, he said. “For us, the problem is that the legislation doesn’t give us the power to get the data, so how do we assess the hazards of the chemicals which are used in commerce and to which there is exposure of the human population when there is no information on their toxicity?” he asked.

When the agency investigated whether there were new possible solutions to this problem, they recognized the potential of the high-throughput screening “and we blatantly stole from pharmaceutical drug discovery the high-throughput screening assay approach,” Kavlock said.

The agency’s ToxCast™ research program was its first effort to run 1000s of untested chemicals through the high-throughput screening assays. “Our unique contribution is that we’re running 100s of assays on these chemicals and trying to recover the biology that leads to [toxic] events,” he said.

The ToxCast™ program’s goals are to identify targets or pathways linked to toxicity, and to obtain assays for these targets or pathways. They are using the assays to screen large numbers of chemicals, starting with those for which there is a lot of toxicological information. Next, we began to develop predictive *in vitro* and *in vivo* models and use the signals they identified to prioritize chemicals for testing and to design testing strategies, Kavlock said.

When he and other EPA scientists talk to researchers about the agency’s strategy, they get a variety of responses, Kavlock told the audience. Some of these responses relate to microphysiological systems, but others do not. The reactions include:

- The biology is too complicated to be addressed by this reductionist approach.
- The approach may miss toxicities due to emergent properties of cells and tissues.
- New testing approaches don’t include feedback loops that afford resiliency.
- We will never know all of the toxicity pathways.
- The approach does not include the liver.
- The results from assay (x) do not produce the “right” results for a given chemical.
- The results from assay (x) disagree with the results from assay (y).
- Some chemicals of interest have properties, such as their volatility or solubility, that don’t allow them to be tested.
- Everything is likely to be tagged as hazardous because it generates a positive *in vitro* response.
- The approach doesn’t consider the dose-response relationship.
- How can we be sure about protectiveness for human health?
- Finally someone is trying to tackle the problem, so let’s give them a chance.

Kavlock showed a slide with data on the program’s progress to date. The agency’s phased approach began with testing about 300 chemicals, mostly widely used pesticides, with about 600 assays for about 1100 endpoints. This first phase of the program ended two years ago. The agency is currently finishing up the second phase of testing, which focused on evaluating about 1000 chemicals, including 130 compounds that failed in clinical trials that were donated by six pharmaceutical companies. Another 880 compounds were run through 50 assays to provide information on interactions with various components of the endocrine system.

In partnership with the NIH and the National Toxicology Program, EPA is evaluating more than 8000 chemicals through the Tox21 program at the NIH's Chemical Genomics Center, which includes 25 assays per year, initially focused on steroid receptor interactions followed by cellular response pathways.

One of the studies that Kavlock said the agency has found to be valuable for helping interpret the data is "reverse toxicokinetics." Conducted in partnership with Rusty Thomas of the Hamner Institutes, the process involves two assays (a primary hepatocyte clearance assay and plasma binding assay) and software that estimates the oral dose that is required to produce steady state concentrations in the body equivalent to those associated with *in vitro* toxicity. To explain why the reverse toxicokinetics process is useful, Kavlock showed a slide with the half-maximal activating concentration (AC_{50}) for compounds in the *in vitro* ToxCast™ assays. The AC_{50} values for many tested compounds were very similar, around 10-20 µM. In comparison, the distribution of oral equivalent values calculated through the reverse toxicokinetics process for the same chemicals ranged over two orders of magnitude. Researchers are now combining the reverse toxicokinetics process with exposure estimates to drive a risk-based prioritization scheme with data on both hazard and exposure, Kavlock said.

Kavlock said he agreed with Thomas Hartung, when he pointed out earlier in the day that validation is an important consideration when dealing with new technologies such as microphysiological organ systems. You need to define the relevance, reliability and fitness for purpose of new tests, such as these, Kavlock said.

"It's nice to see all of these new technologies that we're beginning to assemble for predicting human toxicity. It's how we integrate all of these systems together and how we share knowledge and share information that will help us to move forward – whether you're interested in drug development or whether you're interested in the effect of chemicals in the environment," he said.

Kavlock concluded by enumerating some of the issues that he feels researchers would do well to consider.

- What can the high-throughput testing systems evaluated through Tox21 contribute to research using multi-tissue platforms?
- Where would these tools best fit into the environmental chemical hazard assessment paradigm?
- What level of complexity do we need to improve predictive models – two organs? All ten? Will adding more organs create more challenges than benefits?
- What steps are needed to ensure chemical identity, purity, stability and distribution?
- What are the physical-chemical properties – such as water solubility, log P, molecular weight, and vapor pressure – that are amenable to multi-tissue platforms?
- What steps are needed to characterize the metabolic potential of the systems?
- As we move toward personalized medicine, to what extent do gender differences and genetic variability need to be accounted for?
- What are the key proofs of concept that must be demonstrated to satisfy regulatory agencies and bodies?

The final speaker at the conference was Melvin Andersen, the Charles E. Hamner Distinguished Fellow at the Hamner Institute. His talk asked the audience to ponder whether everyone agreed about what they wanted to achieve with the microphysiological platforms being developed.

At the outset of his talk, Andersen stressed that he was not an expert on microphysiological systems, but rather on toxicology, which he has been practicing for 42 years. He focused his remarks on his impressions of the systems and their potential to aid in assisting in the regulation of environmental agents. He said he has a whole series of unsolved problems that he would like to see addressed.

Is this a "build it and they will come" situation, Andersen asked? It is clear that scientists can build a microphysiological platform to integrate a group of disparate tissues and cell types to evaluate the safety of chemical compounds, he said. These tools give us a lot of opportunities with different aspects of toxicity testing for probing mechanisms, he continued. To Andersen, this potential raises the question about what the new systems do that current systems are incapable of achieving as well as where the successes achieved at this early stage of development will lead us.

One idea is that the new platform will enable scientists to test more environmental compounds more efficiently. This meets a need identified in the National Research Council's *Toxicity Testing in the 21st Century* report, Andersen said. Perhaps the platform will also help speed the testing of new drugs and biological products when human efficacy studies are neither ethical nor feasible that currently fall under the FDA's animal rule by allowing studies to be conducted on a limited human platform.

But how do scientists know that the new platforms will be helpful in this regard, Andersen asked. In order for the platforms to have value, scientists need to be confident that they are capturing the likely toxic responses, Andersen stressed. A key issue that Andersen said he would like to see addressed is how microphysiological systems can be used mechanistically with human tissue aggregates to get a better understanding of modes and mechanisms of action. "But in order to get somewhere, we have to know where we're going to go," he said.

"No matter how we hook these tissues together, they don't add up to a human-on-a-chip. Each one of them lacks critical components to be a full tissue, but it still can be tremendously useful. Eventually, the conversation has to be focused on a restricted platform to validate the *in vitro* findings, he said.

"I think that we have to be careful about creating hype when we say 'human-on-a-chip,'" Andersen said. "The challenge is to avoid getting into a situation where we get overcome by events because of unrealistic expectations that we have for the technology."

Andersen suggested that the scientists developing microphysiological systems consider what design criteria needs to be developed to insure that "tissue exposures" on the multi-tissue test bed are representative of the exposures that are expected *in vivo* inside test animals and humans.

He also asked the assembled scientists to consider what minimal characteristics are needed to have a realistic blood plasma volume for connecting the tissues in the systems. Similarly, he asked how the portions of the systems involved in clearing toxicants from the human on the chip system will be designed to provide relevant plasma concentration time curve (AUC) measures. “We have to get the dose right for the tissues before we can ask the question of are we getting the responses right,” he continued.

Another question Andersen asked the audience to ponder was what key test compounds should serve as prototype chemicals for “validating” the systems. Prototypes need to include compounds that have direct effects on specific tissues and others where metabolites formed in one tissue, usually the liver, go on to damage a second tissue connected through the circulatory system.

Andersen’s professional focus on the pharmacokinetics of how the body responds to exposure to drugs and environmental contaminants inspired him to contemplate the kinetics underlying how microphysiological systems function. To him, this is a matter of how the system plumbing is designed, not just the decision of what tissue to include on the platform. For example, he asked the audience “how do we assure that there is a sufficient amount of liver tissue to produce metabolite tissue exposures in the circulating media that are representative of the concentration we would see *in vivo*?” In other words, is the volume of the liver in the microphysiological system correct in relationship to the total volume of the system, he explained.

Another question is how the scientists create plasma filtration modules to function as part of the system. “Does the kidney-on-a-chip allow for filtration of Phase 2 metabolites, or do you have to have some other filtration assist device to maintain realistic concentrations of test compound and metabolites for extended periods of time?” he asked.

Andersen also raised issues related to how the dose added to the system distributes to the various tissues. For example, how can the systems simulate exposures via oral, intravenous, inhalation or dermal routes? In addition, do the current designs account for the normal relationships of tissues with both parallel and sequential arrangement of tissues within the circulatory system?

If the goal is to improve evaluation of the effects of chemical exposures, Andersen suggested that it may make sense to consider more carefully which tissue systems should be included in the first test systems. Rather than rushing to combine all of the organ systems to produce a ‘truncated human-on-a-chip,’ Andersen pointed out that the collaborators may do well to show that a more limited platform can faithfully represent expected tissue exposures expected in an intact organism.

8 Final group discussion

After Dr Andersen’s talk, the conference’s participants were invited to come up to the podium with comments, observations and responses to all of the day’s talks.

Ingber noted that John Wikswo at Vanderbilt recently published a critical review on scaling and systems biology for integrating multiple organs-on-a-chip (Wikswo et al., 2013). The article raises many of the same points as Andersen did, Ingber said. “It’s critical to delineate the problems because people will come up with solutions,” he said. “I don’t think we have the solutions right now… but I can think of ways of accomplishing what we need to do. The important point is that people realize that there’s this level of complexity to it.” There are going to be problems, and it’s important to balance the need to get people excited about this and overhyping it.

Andersen commented that leveraging biological design principles is as true for the circulatory system as it is for the individual components being developed to go on chips.

Taylor contended that “one of the remarkable things about this program is that it is ‘DARPAesque,’ in that challenges are made that are impossible to do from the traditional government funding grant.” These projects wouldn’t get funded through the NIH’s traditional R01 grant, he said. “We’ve become very focused and experienced on incremental science.” DARPA has a strong track record for achieving the impossible, including allowing inter-computer communication via the DARPAnet/internet and building radar-evading airplanes, he pointed out. “All of the work that’s going on which will have no practical impact over the next five years will lead the way to fairly important things,” he predicted. “Along the way, some targeted focused activities can occur,” he added.

Andersen agreed with Taylor that the program is DARPAesque. “In the last five or six years, Bob and I have been trying to bring the community along to doing business differently. And that community is not DARPAesque – it has concrete on its feet,” Andersen said. “And how you join this kind of rapid movement of new ideas with important regulatory needs to get people involved in a dialog requires a lot of effort,” he continued. He expressed his belief that the community’s conservative nature is inspired by the need to be safe. This demands that new tools be validated in some way.

Stevens commented that he was surprised by the pragmatic nature of much that was discussed. Given that the new tools can’t solve all problems at once, he asked where the participants saw the greatest opportunity overlapping with the greatest need.

Taylor responded by saying that he believed that combining disease models with iPS cells would enable researchers to look at a broad array of backgrounds and take advantage of the complexity of the 3-D models that generate functional activities in situations where you do not see the signal.

Stevens agreed. “I think if we had a 20% improvement in our ability to predict efficacy we would have a much larger impact on overall productivity and result in a 50% reduction in terminations due to clinical safety,” he said, stressing that this wasn’t intended to be a value judgment. “It’s just a factual statement based on the numbers.”

However, Stevens expressed the opinion that there may be some value in putting animal cells-on-a-chip. “The idea is that we’re going to do this in a way that actually shows a species difference,” he said. How do we know that putting human cells-

on-a-chip is a better model if we haven't actually asked if it produces a species difference, he asked. As an example, he pointed to what he termed some low-hanging fruit for comparing results. We know peroxisome proliferation doesn't work in a canine liver culture, but it does work in the rat, he said. But there are more complex systems where we have species differences but we really have not attempted to replicate the differences. There's an opportunity on the efficacy side in that we know that some models handle lipids much differently than humans do, Stevens noted. Replicating these differences could be a way to demonstrate the utility of the microphysiological models, he said.

Taylor responded by noting that some project funding requirements necessitate the use of human cells and tissues. Along the way, we're looking at rodent models to work out the systems biology, he said. He said it is crucial to have a rodent-on-a-chip to be able to identify rodent toxicity pathways. He predicted that this is something that will occur as systems biologists take up the technology.

Ingber pointed out that the research done to date using the microphysiological models with human tissues has generated very interesting data on disease models, such as the virus propagation of disease models. They have also generated interesting data on absorption and drug efficacy. He predicted that the use of animals in toxicity testing may fall by the wayside.

Kavlock commented that one of the criticisms of the ToxCast™ program is its focus on animal toxicity and the notion that it is not the best "gold standard." About five years ago, he gave a poster on the topic at a Society of Toxicology meeting, he said. A pharmaceutical toxicologist at Pfizer who read the poster asked Kavlock what he could do to help. In response, Kavlock said "give us your failed drugs." Pfizer delivered more than 60 drugs that had failed in clinical trials due to renal, liver and cardiac toxicity. Testing the drugs proved a way "to give us an anchor to the human population," he noted.

After that, Kavlock continued, Mike Holsapple of Health and Environmental Sciences Institute *(*HESI*)* suggested that pharmaceutical companies donate additional failed chemicals. Eventually, five other companies also provided failed drugs to EPA's program. The drugs helped prove the program's usefulness, Kavlock said. FDA should be in a position to convene the pharmaceutical industry to engage in a community resources effort to create chemical pools to use for testing the microphysiological systems.

Thomas Hartung pointed out the contrast between the discussions related to how microphysiological systems can be used in drug testing and their use for evaluating environmental chemicals. "The problems are completely different," he said. In the case of drugs, the goal is to characterize one substance extensively with a focus more on efficacy than toxicity. In comparison, regulators would be happy to have any information on many environmental chemicals.

Despite their differences, Hartung said he believes it is important for representatives of the two sides to communicate with each other. A subject that has not received as much attention is the high-content approach to both imaging and 'omics. This is technology which can be used to identify mechanisms," he said. In the case of microphysiological systems, Hartung argued that it is the "only way to really make optimal use of very scarce resources."

"We will not have an endless number of these models to use in parallel to study all of the concentration responses in all of the different substances we want to test," Hartung continued. It is also possible to collect meaningful data by applying metabolomics and transcriptomics on these types of assays. "I think that we should keep in mind that these two extremes have a middle ground which is high-content omics technologies helping us with the identification of mechanisms and getting all of the things together under one umbrella," Hartung said.

Myrtle Davis of NIH also commented on the species toxicity question. A good example of a question that was asked of several agents at the National Cancer Institute was whether or not the toxicities to bone marrow seen in dogs, rats and mice from these agents would also be seen in humans. "In the phase I clinical trial, this was a critical question," she said. What ended up happening was that the bone marrow assay was amenable to a comparison between human bone marrow, and mouse, dog, monkey and rat, as well as any other species from which the cells could be derived. "What we were able to do was to get the sensitivity read in order to figure out what the first-in-human dose should look like. For example, the researchers needed to decide if the dose should it be 5 or 10 times above what they saw in the preclinical setting, and to determine if they would be putting patients at undue risk if they used the straight efficacy model.

"I think that there is room right now for these models to inform whether or not a species sensitivity is there," Davis said. "Species specificity is a different question. It's a higher bar." For sensitivity, the bar is much lower, and researchers could evaluate, for example, whether the rat system in the same context is more sensitive to the toxicity as the human system put together in the same context. "Those are answers that could be very helpful for decision-making when it comes to the first-in-human dose," Davis said.

James Hickman of the University of Central Florida pointed out that this is not a brand-new field. It's been around for about 15 or 20 years, he said. A *Nature* article first coined the "body-on-a-chip" terminology in 2006. A researcher at Cornell University has worked out and published data on the volumes, kinetics, and flow rates for laboratories-on-a-chip. This researcher has shown with tumors in livers on chips that an active metabolite from a cancer drug causes the tumor as opposed to the drug itself. "So we don't really have to re-invent the wheel. Let's make sure that we're not going through from the very beginning because there is a whole series of publications out there," Hickman said.

Andersen responded that he has collaborated with the Cornell researcher. "I think the tools are here, the engineering tools and the biology. It's just getting them to be appreciated." He also pointed out that the *Toxicity Testing in the 21st Century* report discussed new methods that would do testing in cell lines, primary cells, stem cells and tissue aggregates. From his point of view, the piece of the work that is the most interesting for future testing is the tissue aggregates on the chips. It's not quite a whole tissue, but important pieces. "These can become more standardized assays that integrate function and look at human cells and can probe particular modes of action depending upon the tissues. I might sound a bit negative about the human-on-a-chip, but these tissues-on-a-chip are a step between doing high-throughput work and getting assays that are going to let us move on with a

bit more confidence toward doing dose-response and other evaluations of safety, especially in the field of environmental chemicals," he said.

Lewis Kinter, Senior Director at AstraZeneca, responded to Stevens' earlier request for suggestions for low-hanging fruit. His vote was for a proximal tubule organ as a platform for trying to get to the mechanisms of how compounds that can harm the kidneys, such as cisplatin, are actually damaging a segment of the nephron. "We're well aware of the toxicity of these agents but we're no closer to understanding what we should be designing away from to avoid that toxicity," he explained.

Kavlock said that Ingber's descriptions of his group's work with the proximal airway could prove very interesting for his agency. It might help identify the effects of small particulate matter, which represents a significant public health issue in the U.S., and perhaps even the mechanisms underlying these effects. The agency is currently debating whether the national standards for particular matter 2.5 microns in size or smaller ($PM_{2.5}$) in ambient air will be 15, 14, 13, or 12 milligram per kilogram. "The impact is significant because increases in PM are associated with increases in mortality in sensitive subpopulations in this country," he said.

Davis expressed her opinion that bone marrow represents a unique opportunity because there is fairly rich data with respect to bone marrow toxicity. Scientists know what concentrations cause it, what group is reversible, and which patients responded to the drug in terms of efficacy. In addition, researchers have the opportunity to look preclinically at bone marrow toxicity and there are validated *in vitro* assays using human donor cells. "Bone marrow would be a fairly low-hanging fruit with respect to seeing whether these systems are better than what we can do in the clinical setting," she said.

Susan Fitzpatrick of the FDA closed the meeting by thanking all of the speakers and all of the participants, especially those who contributed thoughtful comments. "In putting together this program with the FDA and CAAT, we wanted to create a community of people who were all working toward a common purpose," she said. She credited the National Research Council's *Toxicity in the 21st Century* report for helping to catalyze the creation of such a community. "To move innovation, we really need a whole community of people giving comments like this back to us as we try to look at these new toxicology models," she concluded.

References

Arrowsmith, J. (2011). Trial watch: Phase II failures: 2008-2010. *Nat Rev Drug Discov 10*, 328-329. http://dx.doi.org/10.1038/nrd3439

Babiarz, J. E., Ravon, M., Sridhar, S. et al. (2012). Determination of the human cardiomyocyte mRNA and miRNA differentiation network by fine-scale profiling. *Stem Cells Dev 21*, 1956-1965. http://dx.doi.org/10.1089/scd.2011.0357

Coecke, S., Balls, M., Bowe, G. et al. (2005). Guidance on good cell culture practice. *Altern Lab Anim 33*, 261-287.

Hackam, D. G. and Redelmeier, D. A. (2006). Translation of research evidence from animals to humans. *JAMA 296*, 1731-1732. http://dx.doi.org/10.1001/jama.296.14.1731

Harrison, R. (1910). The outgrowth of the nerve fiber as a mode of protoplasmic movement. *J Exp Zool 9*, 787-846. http://dx.doi.org/10.1002/jez.1400090405

Hoffmann, S. and Hartung, T. (2006). Towards an evidence-based toxicology. *Human Exp Toxicol 25*, 497-513.

Huh, D., Leslie, D. C., Matthews, B. D. et al. (2012). A human disease model of drug toxicity – induced pulmonary edema in a lung-on-a-chip microdevice. *Sci Trans Med 4*, 159ra147. http://dx.doi.org/10.1126/scitranslmed.3004249

Huh, D., Matthews, B. D., Mammoto, A. et al. (2010). Reconstituting organ-level lung functions on a chip. *Science 328*, 1662-1668. http://dx.doi.org/10.1126/science.1188302

Jang, K. J., Mehr, A. P., Hamilton, G. A. et al. (2013). Human kidney proximal tubule-on-a-chip for drug transport and nephrotoxicity assessment. *Integr Biol (Camb) 5*, 1119-1129. http://dx.doi.org/10.1039/c3ib40049b

Kim, H. J. and Ingber, D. E. (2013). Gut-on-a-chip microenvironment induces human intestinal cells to undergo villus differentiation. *Integr Biol (Camb) 5*, 1130-1140. http://dx.doi.org/10.1039/c3ib40126j

Kim, H. J., Huh, D., Hamilton, G. and Ingber, D. E. (2012). Human gut-on-a-chip inhabited by microbial flora that experiences intestinal peristalsis-like motions and flow. *Lab Chip 12*, 2165-2174. http://dx.doi.org/10.1039/C2LC40074J

Kola, I. and Landis, J. (2004). Can the pharmaceutical industry reduce attrition rates? *Nat Rev Drug Discov 3*, 711-716. http://dx.doi.org/10.1038/nrd1470

Olson, H., Betton, G., Robinson, D. et al. (2000). Concordance of the toxicity of pharmaceuticals in humans and in animals. *Regul Toxicol Pharmacol 32*, 56-67.

Pound, P., Ebrahim, S., Sandercock, P. et al. (2004). Where is the evidence that animal research benefits humans? *BMJ 328*, 514-517. http://dx.doi.org/10.1136/bmj.328.7438.514

Rice, J. (2012). Animal models: Not close enough. *Nature 484*, S9. http://dx.doi.org/10.1038/nature11102

Scannell, J. W., Blanckley, A., Boldon, H. and Warrington, B. (2012). Diagnosing the decline in pharmaceutical R&D efficiency. *Nat Rev Drug Discov 11*, 191-200. http://dx.doi.org/10.1038/nrd3681

Seok, J., Warren, H. S., Cuenca, A. G. et al. (2013). Genomic responses in mouse models poorly mimic human inflammatory diseases. *PNAS 110*, 3507-3512. http://dx.doi.org/10.1073/pnas.1222878110

Takahashi, K., Tanabe, K. Ohnuki, M. et al. (2007). Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell 131*, 861-872. http://dx.doi.org/10.1016/j.cell.2007.11.019

Takahashi, K. and Yamanaka, S. (2006). Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell 126*, 663-676. http://dx.doi.org/10.1016/j.cell.2006.07.024

Wikswo, J. P., Curtis, E. L., Eagleton, Z. E. et al. (2013). Critical review: Scaling and systems biology for integrating multiple organs-on-a-chip. *Lab Chip 13*, 3495-3511. http://dx.doi.org/10.1039/C3LC50243K
Yu, J., Vodyanik, M. A., Smuga-Otto, K. et al. (2007). Induced pluripotent stem cell lines derived from human somatic cells. *Science 318*, 1917-1920. http://dx.doi.org/10.1126/science.1151526

Correspondence to
Kellyn Betts
8023 Glenside Drive
Takoma Park, MD 20912
USA
Phone: +1 202 321 6678
e-mail: k_betts@nasw.org