Summary of the July 2010 Workshop, "21st Century Validation Strategies for 21st Century Tools"

Invited responses and discussion

Session 1: Integrated Testing Strategies

Plenary Speaker **Joanna Jaworska** of Procter & Gamble-Europe delivered a paper on "Integrated Testing Strategies (ITS) – Opportunities to better use existing data and guide future testing in toxicology" (see page 231).

Rick Becker of the American Chemistry Council complimented Jaworska on her thorough analysis of the issues and discussed the challenge of integrating the new model and methodologies proposed by Jaworska into risk assessment in the United States. Becker proposed using the Organization for Economic Development (OECD) principles and guidance for Qualitative Structure-Activity Relationships (QSAR) as a model for that process. Integrated Testing Strategies should utilize relevant information from multiple sources, he said, but at present, ITS should supplement the current risk assessment model rather than replace it. It is not yet clear, Becker said, whether or not ITS can replace specific sequential testing batteries. However, while moving forward, it should be possible to use ITS in initial hazard classification to prioritize, using animal models only when Tier 1 data indicate the need for further testing. He proposed using the same endpoints as the OECD for Tier 1 testing – acute toxicity, skin and eye irritation, dermal sensitization, and genotoxicity. The current challenge is to develop *in vitro* and *in silico* approaches that add to the range and complexity of biological responses necessary for decision-making purposes. Becker said that at present an opportunity exists to marry new technologies, where scientifically supportable, to animal testing.

John R. (Jack) Fowle III, Deputy Director of the Health Effects Division in the Environmental Protection Agency's (EPA) Office of Pesticide Programs, said that Jaworska's presentation was consistent with EPA's current approach. Nonetheless, ITS approaches have shortcomings due to existing knowledge gaps. He too mentioned OECD and WHO principles as models for incorporating new approaches such as ITS into U.S. risk assessment. Current models are inadequate to meet the need to assess the possible hazards of between 25,000 and 60,000 chemicals in use for which no adequate safety information is available, he pointed out. EPA is trying to transition to a more efficient and cost-effective testing strategy. New tests should be hypothesis-driven, transparent and accessible, he said, and must be proven to be equally informative as old tests. A combination of approaches will be required for some time, he said. The gold standard (animal) tests have been useful in the past, he noted, but are not sufficient to address current goals – to devise a better way to prioritize and screen chemicals, to develop diagnostic biomarkers, and to more effectively assess and manage existing (limited) chemical data. EPA would like better ways to determine what to test, when to test and how to test, to reduce uncertainties and move toward hypothesis and mechanism-driven approaches that focus on the effects most relevant to risk assessment and risk management. Achieving a paradigm shift of the sort outlined in the NAS report will require substantial effort over many years. He suggested picking a few key tests to pinpoint an outcome of concern. An example is the ToxCast approach which took 320 compounds of interest, mainly pesticides, screened them by chemical structure, and linked their mechanism of action to various disease endpoints.

Troy Seidle of the Humane Society International was to have delivered the third response to Dr. Jaworska's paper but was unable to attend the meeting. His colleague, Martin Stephens, initiated the discussion period by commenting on a metaphor Dr. Fowle used in his response – that of two houses, one (the old testing strategy) being torn down and another (the new strategy) being built simultaneously. He pointed out that one might also use a different metaphor to describe the process currently underway – building a futuristic car (new testing strategies) using the parts from an old car (animal testing) and suggested that this might not be the optimum way to proceed.

Rodger Curren of the Institute for In Vitro Sciences said that one of the difficulties of implementing the new vision is to educate and train scientists in new methods and approaches. Fowle admitted that this is true, pointing out that most EPA scientists were trained in the last century. Vicki Dellarco of EPA noted that EPA is currently setting up a training program in computational science for its staff and that EPA is looking at high-throughput systems to prioritize compounds for further testing. At some point, some of those approaches might be used in Tier 1 testing, she said.

Another member of the audience wondered how EPA defined "prioritization" of chemicals and suggested that with tens of thousands of chemicals untested, it might be best to use limited resources where they would have the biggest impact, namely

biologically-based, hypothesis-driven tests. Becker said that tier-testing and prioritization are basically the same thing and that *in vitro* tests are used to rule chemicals in or out in an evidence-based way.

A panel composed of all the invited speakers offered further elaboration of these points. Jaworska first pointed out the need for all stakeholders – industry, regulatory bodies and academic toxicologists – to work together in a collaborative way to meet the challenges posed by transitioning to the new approach. She suggested that in his response to her presentation, Becker was reverting to the more heuristic approach characteristic of 20th century toxicology, which she tried to reduce in the strategy she proposed. Becker responded that he views the current mixed approach as a purely transitional one. He suggested a stepwise approach that builds on what is already known to build confidence in new strategies. A construct has been in place for the past fifty years, he pointed out, one that inspires both confidence and inertia. A priority setting approach that is risk or hazard based will permit a focus on chemicals of concern.

Fowle reiterated the need to view the current phase as a transitional one. He pointed out that science is unable to ensure safety because safety is a societal construct. When a chemical is clearly causing damage to a (whole) animal, regulators know it is not safe. But the new approach, which relies on theoretical risk, is difficult to sell to legislators and their constituents. Thomas Hartung of CAAT said that the very advanced approach proposed by Jaworska might be better characterized as a 22nd century approach. Daniel Dietrich of the University of Konstanz said that though new 21st century tools have been developed, they are being used with animal tests based on 19th and 20th century science.

In answer to a question from an audience member about validation, Jaworska said that she attempted to avoid that term in her presentation because the theory (ITS) is based on the laws of logic and that there is no need to formally validate such a strategy. Becker countered that it is necessary to build scientific confidence in the tools that would be used to make specific (regulatory) decisions. He suggested bringing in the new methodologies and using them in an appropriate manner. Another audience member said that although some very elegant new tools are being rolled out at present, in the U.S. regulators appear reluctant to make risk assessment decisions based on those tools. EPA is striving for enhanced interpretation of data, Fowle commented, pointing out that it is never possible to have a complete answer to the question of safety.

Melvin Andersen of The Hamner Institute said that though toxicologist have grown comfortable with certain reference tests, many don't feel comfortable inferring low dose effects from the high dose effects observed in those tests. Vicki Dellarco of EPA said that to reach the vision laid out in the NRC report requires a scientific foundation that still needs to be built and that it would take a substantial investment of both time and resources to build the knowledge necessary to transition to the NRC vision. Andersen said that the process would move faster if there were a consensus on the ultimate goal. Fowle added that though the agency's initial focus had been on screening for prioritization purposes, EPA had never lost sight of the long-term goal, enhanced interpretation of data.

Becker said that there are certain predictive models that can be used at present with greater confidence. A member of the audience asked how much confidence people had in the old animal tools, pointing out that they are full of problems but that people have simply become accustomed to using them.

Thomas Hartung of CAAT closed the discussion by noting that we still don't know what an Integrated Testing Strategy would look like – that it is not simply a fixed combination of tests. It is not a linear process. It is not simple tier-testing but more of an artificial intelligence approach. The (U.S.) regulatory community has shown itself willing to embrace some of this new strategy.

Session 2: Enhancing modern technologies for Risk Assessment

Plenary speakers **Kim Boekelheide** of Brown University and **Melvin Andersen** of the Hamner Institute for Health Sciences presented a paper titled "A mechanistic redefinition of adverse effects – a key step in the toxicity testing paradigm shift" (see page 243).

Raymond Tice, chief of the National Toxicology Program's biomolecular screening branch, delivered the first invited response. Tice noted the complexity of cellular pathways and the difficulties of pinpointing specific adverse effects that might result from perturbations to those pathways *in vitro*. He further noted that it would be useful to identify pathway perturbations that would best permit a redefinition of "adverse effect." With respect to those who advocate replacing animal data with human data, he questioned the extent to which such data currently exists and is available. Human data will be difficult to come by, he said, and thus it will be challenging to move away from animal tests in the near future.

Tice pointed out the challenges of extrapolating between single cell events and biological processes like organ failure. He said that the NTP current testing paradigm has changed over the past few years with a move toward studies providing more mechanistic information. NTP is also trying to move forward in terms of interpreting data in the context of human relevance, he said. In addition, NTP plans to make their standard toxicological data transparent via a mineable database. He suggested that one of the problems encountered in this transitional phase is not knowing what we don't know – in other words, we don't yet know what information about pathway dynamics is relevant for regulatory decision-making.

Tice next discussed the U.S. Tox21 project, which has set the following goals: to prioritize chemicals for more extensive toxicological evaluation, to identify mechanisms of chemically-induced biological activity, to characterize pathways of toxicity in order to facilitate cross-species extrapolation and provide input to develop models for low-dose extrapolation, and to develop predictive models for adverse effects in humans. He noted that federal organizations involved in Tox21 – EPA, NIEHS/NTP, the NIH Genome Research Institute/NIH Chemical Genomics Center (NCGC), and soon, the FDA – are actively developing

a strategy for achieving the goals of Tox21, but that the process is slow given the lack of knowledge about the relationship between perturbations and adverse reactions. Tice pointed out that only twenty-two percent of 1100 human cellular pathways and 35 percent of human disease pathways are currently covered by assays used by the NGGC.

He then posed a series of questions. To what toxicity endpoint should data from *in vitro* assays be linked? (At present, the most likely is disease in animal models). When is a pathway perturbation to be labeled adverse? Which cell types are most useful in *in vitro* studies? Even if the focus is on human cells, which type and from which human? Considering that a major limitation of most *in vitro* assays is the lack of xenobiotic metabolism, how do we incorporate such metabolism into these assays? Finally, given that *in vitro* approaches will result in probabilities of adversity rather than yes-no answers, how will a probabilistic model be used in risk assessment?

Jos Kleinjans of Maastricht University in the Netherlands delivered the second invited response. He said that adverse effect was generally defined by that country's regulatory community as the chance of an adverse health effect in an exposed population within a given timeframe, and that specific definitions of "adverse effect" relied on expert judgment. He pointed out that current testing strategies have limitations, such as a high rate of false positives and inadequate data for assessing human risks from rodent bioassays, yet for thirty years these tests have remained unchanged. Perhaps it is time to change them, he suggested. He pointed to the potential of toxicogenomic-based screens for toxic class prediction and hazard identification and said that it has been abundantly proven over the years that such an approach can distinguish between genotoxic carcinogens and non-genotoxic carcinogens with ninety percent accuracy.

Kleinjans agreed that it is hard work trying to absorb all the new technologies currently being developed into new testing strategies but that they might allow us to move beyond animal testing. He suggested that recent findings in epigenetics signaled the end of the linear non-threshold model. Epigenetic changes are potentially more damaging because they affect multiple genetic loci, he said, while also tending to affect a high proportion of cells exposed to toxins. He also noted that epigenetic inheritance of disease effects continues for decades after initial exposure (using endocrine disruptors as a model). He pointed out that DNA methylation should be a key area of study as there is a growing body of evidence to suggest that exposure to chemicals can induce changes in DNA methylation patterns.

Kleinjans said that he is a believer in the new technologies because they have the potential to provide a better understanding of perturbations in molecular pathways underlying toxic modes of action. Moreover, this understanding can be derived from relevant human cellular models. The mechanistic information needs to be connected to molecular information on human diseases by means of human translational research. Do the new technologies need to be validated against gold standard rodent bioassays, he asked. Scientifically no, but politically probably yes, he said, citing "conservatism" as the reason why it might be necessary to validate *in vitro* assays against rodent *in vivo* tests as a first step. It might be a waste of both animals and money but still politically necessary.

Russell Thomas of the Hamner Institute delivered a final response to Boekelheide and Andersen. He said that they had provided a much needed extension of the approach outlined in the NRC report. The success of the approach will depend on the development of relevant assays, quantitative and qualitative recapitulation of the key network elements in the *in vitro* models, or at least an understanding of when and why they don't work, as well as experimental identification and characterization of key network elements. Relevant pathways have been largely identified, he observed, but how they connect to networks has not. Results must also be placed into the context of what is normal in the whole organism, he said. This will require more than positive controls. The bounds of normal will vary by pathway and network, and therefore the amount of change considered adverse will also shift from pathway to pathway and network to network.

In vitro assays rarely replicate both the magnitude and entire spectrum of *in vivo* response (to toxicants) with complete fidelity, Thomas pointed out. When using prototype chemicals to validate these assays, we face the question of how good is good enough, in both qualitative and quantitative terms.

Boekelheide and Andersen responded briefly to issues raised by Tice, Kleinjans and Thomas. At present, we are working with the wrong model (high-dose toxicity) in many ways, Andersen said. Therefore, is it really necessary to completely replicate *in vivo* high-dose response *in vitro*? No. We are in a dangerous situation right now, he said, in acquiring mechanistic information about high-dose exposures in rats, not humans. The real question that needs to be addressed is how to do toxicity testing to infer safe regions of exposure in humans. Another challenge is how to characterize the relationship between responses to pharmaceuticals *versus* environmental exposure. Boekelheide said we don't need to know everything, only what's important. We don't define adversity very well now, he pointed out – and so we don't have to do a whole lot better to improve on current models.

Andersen said that, regardless of whether the model under discussion is *in vitro* or *in vivo*, we are looking at an operational model of adversity. With chemicals where we have clear modes of action, there can be agreement that the adverse effect observed *in vitro* is sufficient. Boekelheide mentioned that seventy percent is the usual margin of sensitivity. He proposed doing better than that with the new approach. An audience member pointed out that it would be impossible to address all potential pathways of toxicity with a single test. Tice said that multiple pathways intersect in multiple ways and that, if you don't recognize that fact, you can't identify an adverse effect. Kleinjans noted that an advantage of modern tests is that they generate a wealth of information; the challenge is interpreting all that data. But progress is being made in that area too, he said. The goal is deriving a mechanism of action, and it doesn't matter how many pathways are involved. One hundred percent certainty is impossible, he pointed out.

A member of the audience pointed out that there will be a level of uncertainty for any new *in vitro* assay for which there

won't be a readily available reference point. If we don't know how they are going to perform, how will we use them to regulate? Andersen said that current tests are not delivering the necessary information. Do we know how we are going to get to the new paradigm, he asked? No. But it is critical to try because the system is in need of change. The key questions ask: how we are going to do it, and how long it is going to take? We have the opportunity to start making progress now, he noted, and we need to take advantage of what we already know by targeting certain systems and starting to work with them.

An audience member said that he was hearing a desire to move away from comparisons to animal data. He wondered whether it might not be possible to make better use of epidemiological data. Kleinjans agreed. He suggested occupational exposure to benzene as an example and proposed imaging studies using human volunteers to test substances that are part of daily life.

With respect to neurotoxicity, an audience member pointed out that there are twenty-four recognizable areas in the brain and at least six different types of cells. How can one practically assess the brain *in vitro*, he asked. Boekelheide said that the pathways are canonical and operate in all cell types. So it may not be necessary to use neurons to assess neurotoxicity. Also, the current study of rats and mice might not provide a very good assessment of human neurotoxicity either. A discussion of how to reduce the "fear level" involved in adopting the new approach followed. Andersen suggested using everything we have at present as we try to figure out a way forward.

Session 3: Evidence-Based Toxicology

Thomas Hartung, director of CAAT, presented a paper titled, "Evidence-based toxicology – the toolbox of validation for the 21st century?" (see page 253)

William Stokes of the U.S. Public Health Service and NICEATM began his response to Hartung's paper by pointing out that validation has been successful in changing methods used for regulatory decision-making. The current challenge is figuring out how to make validation more nimble, flexible and cost-effective. He strongly supports the proposal to move forward with ITS, as single methods don't provide the information needed. He also supports the pathways of toxicology approach spelled out in the NRC report. Regarding the issue of validation, he said it wasn't too long ago that methods were validated by a ring study; more recently optimized and well-designed studies have been used. With respect to Hartung's proposal, he suggested that the evidence-based approach may have its greatest value when multiple studies of specific chemicals are available. He wondered whether EBT is applicable when multiple studies haven't been done.

Stokes pointed out that validation is an ongoing iterative process of considering all available evidence. Retrospective studies are needed; at present most methods are based on prospective studies. He defined scientific validation as the process of determining the usefulness and limitations of a test method/strategy for a specific proposed decision-making purpose. Those purposes may differ, he noted: replacement (of an existing test), screening decisions, prioritization, or acquisition of mechanistic adjunct data. He stressed the need for more reference data to anchor mechanistic pathway data to phenotypic changes in humans wherever possible so that new biomarkers can be integrated into testing strategies.

Anthony Cox of Cox Associates delivered the second response to Hartung's presentation, providing a mathematical definition of validation. The ultimate goal, he pointed out, is better risk management decisions. Many of the technical challenges have been solved already, he said, and newly developed biometric tools are waiting to be used. Real-world dynamics are often non-linear, he observed, noting that the same mechanisms that maintain homeostasis lead to disease states. The same pathway can have the opposite effect based on exposure to a toxicant over time once a tipping point is reached. The question is how large a perturbation will lead to the disease state.

After providing a mathematical model of network dynamics and pathway analysis, he pointed out that pathways assemble in feedback loops and that it is the dynamics of the network (system feedback) that matter. The tools to track this are increasingly well-known in the biometric literature but not well known to those doing risk-assessment. Many of these tools can be purchased off the shelf.

Michael Holsapple of ILSI-HESI, current president of the Society of Toxicology, delivered the final response. He said that we are not even close to having a consensus on what constitutes an adverse effect using the new approach. He noted the importance of transparency in science-based decision-making. He said that a paradigm shift of this magnitude cannot be addressed by traditional validation methods. He hasn't made up his mind yet about the evidence-based approach to validation proposed by Hartung, but he does think it is an appropriate addition to the tool chest if the expert working group process is objective and transparent. Transparency enhances the credibility of data, risk assessment processes, and decision-making, he said, and a lack of credibility at any of these steps compromises the credibility of the entire process.

Raw data for toxicity studies should be more widely available, he said. Challenges include confidentiality and proprietary concerns, the volume of data, and the need for a central database.

Hartung responded to Stokes, Cox and Holsapple by noting that the evidence-based approach is meant to assemble all types of available information – and not just for validation purposes. The complexity of pathway dynamics will be a dramatic problem. Transparency is one of the hallmarks of the evidence-based approach, he said. A member of the audience pointed out that there seemed to be a tension between the evidence-based approach proposed by Hartung, with its emphasis on consensus, and Cox's comments on causal analysis. He wondered how it is possible to move forward in a flexible manner while still providing stability to markets, with industry requiring reassurance that their testing methods will meet the approval of regulators. Hartung replied that a kind of science that was simple enough for most to understand is being replaced by a much more complex approach in which "the computer is the study

director". It is not easy for those who make the decisions to communicate them, and that's why validation is so important in building credibility via transparency. He pointed out that we no longer live in a world of bad chemicals and good chemicals; science cannot offer those judgments. No producer can ever close the books on a particular chemical. We need to do more sophisticated post-marketing evaluations and to declare the uncertainties in judgment so that consumers can make informed decisions.

Stokes noted that, as both the volume and complexity of data produced by the pathway-based approach to toxicology increase, the biometric tools described by Cox will become absolutely necessary to manage the data. Regulatory agencies cannot close the books on any one chemical or class of chemicals for public health reasons, he said. Having all the raw data on the table would be very helpful. Debbie Lander of DuPont Global Health Centers asked how the evidence-based approach differed from a weight-of-evidence approach, and Hartung defined the distinction between the two. Weight-of-evidence offers a shortcut without a strong conclusion where evidence-based medicine applied to toxicology poses a structured question with a strategy to gain the necessary information and how to evaluate it. The entire process is peer-reviewed, he pointed out, and is very rigorous and transparent.

Audience member Paul Locke of Johns Hopkins asked who should do validation. Stokes replied that validation could be done by anyone. ICCVAM then looks at the results of validation studies and makes recommendations based on scientific validity for proposed uses. Other national authorities do the same; it is recommended that those doing validation studies be in touch with the ultimate users, he said. Hartung added that it is important to get as many people involved as possible; OECD guidelines offer performance standards.

Alan Goldberg of CAAT asked Stokes what was required to establish that a compound was non-toxic, and Stokes said that the new terminology was "not required to be labeled as toxic." Every substance has some potential to cause eye or skin irritation, he pointed out. Another audience member said that scientists are sometimes averse to publishing negative or inconclusive results. Holsapple said that people are not averse to publishing those results but it is difficult to find a journal willing to publish them, and editors should be encouraged to publish negative results. At SOT meetings, he pointed out, at least one-third of papers deal with negative results.

Richard Judson of EPA predicted a coming crisis in validation because of all the new information being generated by new methodologies and asked if it would be possible to validate a body of evidence even though the individual assays have not been validated. Hartung said that it would be possible to bundle these things, for example, reporter gene assays. Many will be scientifically validated by proving they identify pathways. Stokes said that there are two types of validation – biological validation of an assay and validation of a standardized protocol. An audience member asked how to resolve the need to validate with the precautionary principle. Hartung said that legislative and societal pressures are very important as an engine for change. Industry is waiting to see what the regulations will be, rather than developing new approaches independently, because it is risky to do something new.

Another audience member commented on the recommendation that raw data be put online for expert review. He raised the topic of conflict of interest on the part of reviewers. Stokes said that NTP convenes numerous panels each year, and sometimes people serve when they have a conflict (cases where individuals have the potential to benefit financially from the product under review), but they can't vote on outcomes. Holsapple pointed out that everyone has biases and that it would be impossible to put together an expert panel that didn't include a few people with very strong biases but that is different from a conflict of interest based on potential financial benefit.

Finally, an audience member said that industry needs a demonstration of how the evidence-based approach would work. Hartung said that at present it is only a vision which needs to achieve critical mass. Every year, ten trillion dollars worth of products are traded, depending on decisions made by people as in this room, he pointed out, so it will take some time to implement the new approach.

Session 4: Endocrine disruption in application

Daniel Dietrich of the University of Konstanz presented a paper titled, "Courage for simplification and imperfection in the 21st century assessment of 'endocrine disruption'." (see page 264)

Chris Borgert of Applied Pharmacology and Toxicology Inc offered the first response. He said that he was in absolute agreement with Dietrich. Endocrine disruption has rarely been demonstrated in humans. Causal relationships between chemical exposure and adverse effects have rarely if ever been established by epidemiological or animal studies. Human *in vitro* and *in silico* data have greater relevance. One consequence of the present approach is extrapolation rather than causation. Extrapolation means the requisite data are missing; it is an admission that we lack an understanding of causal mechanisms. Risk assessment is therefore a misnomer. Often we are measuring one thing as a surrogate for another, he said.

Borgert stressed the need for counter-factual methods, perturbing the system to measure downstream effects by mapping a causal network. These mechanisms and pathways are multidimensional, he pointed out. We now have the tools to use human systems. Extrapolation must be replaced by translation. That requires demonstrating causal relationships between *in vitro* and *in silico* measurements and adverse effects. Counterfactual study designs are required to demonstrate causality. He expressed a wish for advocates of 21st century toxicology to be less defensive and to enthusiastically embrace these challenges in order to realize the full potential of the new technologies.

Richard Judson of EPA delivered the second response to Dietrich's paper. He said that though there was much public concern about EDCs there was little human data to support it. High-dose animal studies implicate some chemicals; human effects will come from strong interactors. Much of the *in vitro* screening uses only animal tests, he noted. EPA's approach involves

prioritization, dose relevancy and investigating and broadening the range of phenotypes of concern. He mentioned ToxCast Phase 1 studies, in which of 309 chemical structures tested, the EPA also required 64 to go through extensive ED animal studies. By the end of 2011, Tox21 will have tested 10,000 compounds including for endocrine receptor and androgen receptor activity. Any new potent chemicals should be obvious, he said, based on experience testing reference, positive control chemicals. A number of chemicals have tested positive for AR/ER/TR activity in these assays but most are not very potent (any activity only occurs at high concentrations). This reduces concern about the human health effects of those chemicals. However, there is still concern for ecotoxicity because even weak EDCs have been shown to be toxic in aquatic species.

Kate Willett of PETA delivered the final response to Dietrich's paper. She noted the need for a mechanism-based approach to evaluating endocrine active substances. She identified an opportunity to use human data in this field of research. Hormone levels vary naturally within organisms, she pointed out. The same pathway can have different functions in different cell and tissue types. Also, hormone receptor structures and activity can vary dramatically between species. For regulatory purposes there has been a primary focus on sexual and reproductive function; the term endocrine has been left vague or oversimplified.

A new approach incorporating multiple pathway characterization, elucidation of pathway interaction and allowing risk assessment of mixtures is necessary, she said, to deal with the complexity of data generated by new tests. She advocated incorporating mechanistic analysis into existing test methods and feeding all data into databases for system models and pathway models. She also supports immediate application of ITS. Mechanism-based information can be progressively integrated into the existing regulatory framework until animal replacement is feasible, she said, taking us to a future that is both safer and more humane.

Alan Goldberg of CAAT began the discussion by asking if any health effects had been noted in populations that consume large amounts of phytoestrogens in their diets. Judson of EPA noted that one of the hottest chemicals in the recent EPA study was geneistin, but Asian people absorb large amounts of geneistin in their diets, seemingly without any ill effects. Borgert suggested that perhaps Asian people have developed a metabolic capacity to clear it due to their long exposure to geneistin.

A member of the audience asked, if there is nothing to worry about (re: human health effects of EDCs) why are we spending all this money researching them. Dietrich said that spending an hellacious amount of money on a non-issue creates a lack of funding for issues of real importance. He suggested that one of the reasons for continued high levels of funding for EDCs is that some investigators rely upon that funding to support their labs and so are reluctant to say that these studies aren't necessary. He also pointed out that the way that study results are presented to the public is part of the problem.

Another audience member asked Borgert to elaborate on how counter-factual studies can be better utilized to help move forward to the new paradigm. With respect to EDCs, Borgert said that a necessary condition was for the chemical to interact with the estrogen receptor and that condition has not been satisfied by the majority of these compounds. He wondered what would have happened in the von Saal experiments if the mice in the study had also be treated with an estrogen-receptor blocker.

An audience member said that perhaps it would be better to do mechanistic tests earlier in the tier-testing process to facilitate better understanding of the mechanisms of endocrine disruption. Dietrich agreed and said that he saw an opportunity to eliminate the higher tier animal tests entirely. He believes that the amount of money being spent on animal studies for endocrine disruption is political appeasement due to the lack of hard evidence supporting human health effects. He said he is not ready to give up on the idea that these chemicals that hit the ER and AR receptors are completely safe. It is no surprise that these chemicals that cause reproductive fitness issues in rats and mice correlate with ER and AR. Data on EDCs from rodents are not totally relevant to humans, he pointed out.

Andersen said that every concern raised about EDCs in this context could be applied to other types of toxicity testing. For most of his career, he said, he has been faced with the challenge of how to use animal data to predict human health effects. He believes that EDCs should be a case study for using mechanistic information to make better risk management decisions. He said that initially we have to be able to connect the dots across levels of organization.

A discussion about threshold and pass/fail criteria followed. Andersen pointed out that many of the pharmaceuticals we use daily have a very narrow safety index. An audience member asked about chronic exposures over a lifetime and noted that the incidence of prostate cancer in American men is very high versus Asian men – until the latter move here. The only way to figure out the reason is through mechanistic studies, he said. Borgert observed that the public believes that breast cancer is related to total estrogen exposure. A member of the audience said that it is difficult to educate the public about the difference between hazard and risk. If an effect is noted, he said, the public equates that with risk.

Hartung of CAAT ended the discussion by noting that he had learned more than usual from the discussions at this workshop. Clearly, there is a long way to go before we have really integrated testing strategies, he said. However, discussions are underway about next steps, and that was not the case three years ago when the NRC report first came out. He is happy to see that the concept of evidence-based toxicology is resonating; people haven't exactly embraced it, but they are willing to consider it. He was surprised by the lack of controversy which greeted Dietrich's paper. He found that encouraging. He noted that CAAT is committed to promoting this important discussion and helping creating a blueprint for a completely new approach.

Deborah Rudacille