International Journal of Toxicology
32(6) 405-414

DOI: 10.1177/1091581813508436
ijt.sagepub.com
SAGE

Insights and Perspectives on Emerging Inputs to Weight of Evidence Determinations for Food Safety: Workshop Proceedings

Heidi Bialk[1], Craig Llewellyn[2], Alison Kretser[3], Richard Canady[4], Richard Lane[1], and Jeffrey Barach[5]

Abstract

This workshop aimed to elucidate the contribution of computational and emerging in vitro methods to the weight of evidence used by risk assessors in food safety assessments. The following issues were discussed: using in silico and high-throughput screening (HTS) data to confirm the safety of approved food ingredients, applying in silico and HTS data in the process of assessing the safety of a new food ingredient, and utilizing in silico and HTS data in communicating the safety of food ingredients while enhancing the public's trust in the food supply. Perspectives on integrating computational modeling and HTS assays as well as recommendations for optimizing predictive methods for risk assessment were also provided. Given the need to act quickly or proceed cautiously as new data emerge, this workshop also focused on effectively identifying a path forward in communicating in silico and in vitro data.

Keywords

risk assessment, food supply, weight of evidence, high-throughput screening, in silico, computational modeling

Purpose of the Workshop

The International Life Sciences Institute (ILSI) North America Technical Committee on Food and Chemical Safety developed a workshop to gain insight into how computational and emerging in vitro methods could be used to contribute to the weight of evidence that is used by risk assessors in the food sector to determine safety. The practice of food safety assessment has continually evaluated and adopted new methods; however, because of the pace and nature of change in biological effects evaluation methods, it is not clear how some newly developed in vitro data types or computational approaches fit into current approaches to food safety assessment. For example, how would a "cancer gene" activation shown in an in vitro assay at concentrations orders of magnitude higher than could be achieved in foods contribute to our understanding of the likelihood of harm of a commonly used food additive with negative carcinogenicity data? How should we communicate this information to the public while we are trying to understand its significance? How in general can we combine data from these new methods to improve food safety? In particular, it is also not clear where and when these new data types may signal a need for prompt evaluation using toxicity tests used in current practice to evaluate safety. These kinds of questions have also been asked for environmental risk assessment for environmental and occupational exposures to chemicals; however, for food safety assessment, there may be both a need to act quickly in the face of new data and a need to proceed with caution. How do we begin to sort through and prioritize the steps to integrate the new data types into food safety assessment?

To begin to address these issues, the workshop (Table 1) brought together over 70 experts from academia, government, and industry as well as a public interest organization. The workshop began by posing the following questions to participants:

1. How can in silico and high-throughput screening (HTS) data be used to confirm the safety of approved food ingredients?

[1] PepsiCo Inc, Valhalla, NY, USA
[2] The Coca-Cola Company, Atlanta, GA, USA
[3] International Life Sciences Institute North America, Washington, DC, USA
[4] International Life Sciences Institute Research Foundation, Washington, DC, USA
[5] Barach Enterprises LLC, Oakton, VA, USA

Corresponding Author:
Heidi Bialk, PepsiCo Inc, 350 Columbus Ave, Valhalla, NY 10595, USA.
Email: heidi.bialk@pepsico.com

Table 1. Weight of Evidence Workshop Program.

Welcome and Purpose of the Workshop—Heidi Bialk, PhD, PepsiCo Inc, Valhalla, NY
Session 1—Weight of Evidence Inputs: In Silico Data
Computational or Predictive Toxicology and Metabolism: Tools for the Trade? Eugene Ahlborn, PhD, Givaudan Flavors Corporation, Cincinnati, OH
Computation Safety Analysis for Regulatory Decision Making James Rathman, PhD, Ohio State University, Columbus, OH
Mid-Day Presentation—Values, Trust and Science—Building Trust in Today's Food System in an Era of Radical Transparency, Mr. Charlie Arnot, The Center for Food Integrity, Gladstone, MO
Session 2—Weight of Evidence Inputs: In Vitro Data
The Need for a Roadmap for Using Tox21 Methods in Risk Assessment Thomas Hartung, MD, PhD, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD
Risk Assessment Roadmaps and Methods for Using 21st Century Methods Michael L. Dourson, PhD, DABT, ATS, Toxicology Excellence for Risk Assessment, Cincinnati, OH
Evaluation of Food-Relevant Chemicals in Tox21 Agnes Forgacs, ILSI North America 2012 Summer Fellow/Doctoral Candidate, Michigan State University, East Lansing, MI
Weight of Evidence for Targeting "Second Tier Testing" from HTS and EDSP Approaches Michael DeVito, PhD, NIH /NIEHS, Research Triangle Park, NC
Key Events Dose Response Framework (KEDRF) Samuel Cohen, MD, PhD, University of Nebraska Medical Center, Omaha, NE
Session 3—Perspectives on Applying In Silico and In Vitro Data into a Holistic Weight of Evidence Approach
FDA Perspectives on the Application of 21st Century Toxicology Methods Used in Risk Assessment and Regulatory Decision Making David Hattan, PhD, FDA CFSAN, College Park, MD
Perspectives from the Health Environmental Science Institute (HESI) Risk 21 Program Alan Boobis, OBE, PhD, CBiol, FSB, FBTS, Imperial College London, UK
Defining the Elements of a Holistic Weight of Evidence Approach James Bus, PhD, DABT, ATS, The Dow Chemical Company, Midland, MI
Panel Session—Practicality of Weight of Evidence: Where Do We Go From Here?
Moderator: Richard Canady, PhD, ILSI Research Foundation, Washington, DC
Panelists: Roxi Beck, Center for Food Integrity; James Bus, The Dow Chemical Company; Keith Houck, EPA NCCT; Mike DeVito, NIH NIEHS NTP; James Rathman, Ohio State University, Columbus, OH; and Ken Wallace, University of Minnesota Medical School, Minneapolis, MN

Abbreviations: EPA, Environmental Protection Agency; ILSI, International Life Sciences Institute; NCCT, National Center for Computational Toxicology; NIEHS, National Institute of Environmental Health Sciences; NIH, National Institutes of Health; NTP, National Toxicology Platform.

2. How can in silico and HTS data be applied in the process of assessing the safety of a new food ingredient?
3. How would in silico and HTS data help in communicating the safety of food ingredients while enhancing the public's trust in the food supply?

The final session of the workshop culminated in a panel session that not only helped to address these questions but also focused on identifying a path forward in communicating in silico and in vitro data effectively. This workshop is one of the projects by the ILSI North America Committee on Food and Chemical Safety developed in an effort to help fulfill its mission and objectives of promoting a science-based determination of the chemical safety of foods in support of the advancement of public health.

Introduction to Emerging Methods in Predicting Toxicity

Speaker

The Need for a Roadmap for Using Tox21 Methods in Risk Assessment.

Thomas Hartung, MD, PhD, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA

In 2005, the Environmental Protection Agency (EPA), with support from the National Institute of Environmental Health Sciences (NIEHS)/National Toxicology Platform (NTP), requested that the National Research Council (NRC) develop a vision for the future of toxicity testing and a strategic plan to accomplish it.[1] The impetus behind this request was the strong

commitment by the EPA and NTP that future toxicity testing meet evolving regulatory needs in that testing paradigms should (1) readily accommodate the increasingly large numbers of substances in commercial use that are considered to have incomplete toxicity data; (2) incorporate recent advances in molecular toxicology, computational sciences, and information technology; and (3) offer increased efficiency in the design, costs, and extent of animal usage.[2] Two years after the request from the EPA, the NRC released a report entitled *Toxicity Testing in the 21st Century: A Vision and a Strategy*.[1] This strategy was aspirational, nevertheless it described a new approach to toxicity testing that would increase the use of in vitro high-throughput assays and reduce the reliance on (and ultimately replace) animal studies using cells, cell lines, or cellular components to evaluate the effects that chemicals can have on biological processes.[3] The specific near- and long-term goals as described in the NRC report are as follows: (1) to identify patterns of compound-induced biological response in order to characterize disease pathways, facilitate cross-species extrapolation, and model low-dose extrapolation; (2) to prioritize compounds for more extensive toxicological evaluation; and (3) to develop predictive models for biological response in humans.[1] Numerous stakeholders are needed to develop a roadmap for how to use Tox21 methods[4] to support safety assessments. The Doerenkamp-Zbinden Foundation and the Center for Alternatives to Animal Testing (CAAT; http://caat.jhsph.edu) have established the Transatlantic Think Tank of Toxicology (t^4; http://altweb.jhsph.edu/about_us/t4.html). The t^4 was created in part to further the concept of evidence-based toxicology following the role of evidence-based medicine (http://www.ebox.com) and to develop and assess the conceptual needs to enable the change approaches in predictive toxicology, integrated testing, and systems toxicology.[5] So far, 14 workshops have been held and 13 commissioned white papers have been published. In addition, the Human Toxome project, a National Institutes of Health (NIH) Transformative Research grant of 6 partner organizations lead by CAAT, is working to identify the toxicity signatures of substances with a focus on endocrine disruptors (ED). The central goal of the Human Toxome project[6] is to define, experimentally verify, and systematically annotate Pathways of Toxicity (PoT) pioneering this for ED (http://humantoxome.com). The area of ED is well suited for this work, as the physiological pathways of the endocrine system are relatively well understood, making PoT identification simpler for ED than for other potential target organs. This project will also serve to develop a common, community-accessible framework and databases that will enable the toxicology community at large to comprehensively and cooperatively map the Human Toxome combining ''omics'' data with computational models. On a shorter term, integrated testing strategies will enable the combination of different methods to satisfy safety information needs.[7] The following sections of this article describe the current state of the science of emerging, predictive methods in general, and the next steps that are needed to enhance the confidence in these methodologies to support safety assessments.

The Application of Computational Methods in Hazard Identification

Speakers

Computational or Predictive Toxicology and Metabolism: Tools for the Trade?.
Eugene Ahlborn, PhD, Givaudan Flavors Corporation, Cincinnati, OH, USA

Computation Safety Analysis for Regulatory Decision Making.
James Rathman, PhD, Ohio State University, Columbus, OH, USA

Current State of the Science of Computational Methods in Hazard Identification

Computational toxicology is a multifaceted science that incorporates knowledge from a variety of disciplines including but not limited to toxicology, chemistry, computer science, and statistics. The theory behind the use of structure–activity relationships in computational toxicology is that the chemical structure of a compound determines its physical and chemical properties, which in turn determines its biological activity (i.e., toxicology, metabolism, and pharmacology).[8] The initial steps in a safety assessment of a given compound may therefore include predictive computational methods (e.g., in silico techniques) such as structural alerts, quantitative structure–activity relationships (QSAR), and read-across to chemical analogs. QSAR is currently used by the Flavor Extract Manufacturer's Association (FEMA; http://www.femaflavor.org/) and Joint Expert Committee on Food Additives (JECFA; http://www.who.int/foodsafety/chem/jecfa/en/) to support the safety assessments of flavor compounds.

The call to refine and reduce animal testing in addition to the advances in computer technology have enhanced the opportunities to apply in silico techniques in informing science-based health assessments. From an industry perspective, these tools can provide a faster and more accurate evaluation of molecules going through the discovery pipeline, provide additional insights into potential toxicological and metabolic activities, and provide a better understanding of key end points and what testing may be necessary to ensure the safety of evaluated materials. In addition, regulatory authorities around the world are either using or investigating the potential use of computational toxicology in the safety evaluation process.

In order to achieve these objectives, the strengths, weaknesses, and limitations of these predictive methods need to be evaluated. In addition, extensive consideration should be given to how the output provided by computational platforms should be properly applied when conducting a safety/risk assessment. It is critical that these platforms need to undergo

rigorous validation of end points in partnership with the platform providers to continually update and improve the predictability of toxicological and metabolic end points.

Enhancing the Confidence in Computational Methods

Improving computational techniques with respect to the predictability of toxicological end points in addition to metabolic and pharmacokinetic performance are also key to advancing the validation effort. Identifying off-target receptors and substructures for carcinogenic or mutagenic substances would be helpful in improving the predictability of toxicity end points. Pharmacokinetic performance can also be strengthened by understanding how these models can be developed to better describe dose response, metabolism, and the potential for local versus systemic toxicity.

Currently, combining evidence from multiple sources is mainly a qualitative exercise in that multiple outputs are used to determine whether there is a consensus on a given prediction. More quantitative and rigorous statistical methods developed for information theory applications may provide a path forward for quantitatively weighing and combining predictions from multiple sources in order to generate a single predictor for a given toxicological end point (J. Rathman, personal communication, 2013). As an illustrative example relevant to the food and drug industry, the Dempster-Shafer Theory has been used to combine the results from 3 different types of outputs (eg, structural alerts, QSAR models, and read-across to defined chemical analogs) using Ames mutagenicity as the end point of interest.[9-11] Estimating the uncertainty for the final prediction depends upon the reliabilities of the various evidence sources. For QSAR models, prediction accuracies provide obvious measures of reliability; however, for other sources such as structural alerts or read-across, reliability may be best determined by experts in toxicology, including possibly regulatory experts.

Overall, predictive in silico models are based on a wide variety of computational platforms that require ongoing validation. The US Food and Drug Administration (FDA) is nearing the completion of the development of a tool referred to as Chemical Evaluation and Risk Estimation (http://www.accessdata.fda.gov/FDATrack/track-proj?program=cfsan&id=CFSAN-OFAS-Chemical-Evaluation-and-Risk-Estimation-System). This tool will enable the FDA to fully leverage available data generated from modern computational approaches with other data for the pre- and postmarket review of food ingredients and packaging materials.

The Application of High-Throughput Screening Assays in Hazard Identification

Speakers

Weight of Evidence for Targeting "Second Tier Testing" from HTS and EDSP Approaches.

Michael DeVito, PhD, NIH /NIEHS, Research Triangle Park, NC, USA

Evaluation of Food-Relevant Chemicals in Tox21.

Agnes Forgacs, ILSI North America 2012 Summer Fellow/Doctoral Candidate, Michigan State University, East Lansing, MI, USA

Background on Tox21 and HTS

The National Center for Computational Toxicology (NCCT) at the EPA initiated the ToxCast (http://www.epa.gov/ncct/toxcast/) program in 2005 with the initial screening of a phase I library of 320 chemicals across 550 assays. ToxCast is a large-scale experiment using a battery of in vitro HTS assays to develop activity signatures predicting the potential toxicity of environmental chemicals at a cost that is a fraction of that required for full-scale animal testing.[12] Phase I of the Tox21 (http://epa.gov/ncct/Tox21/) quantitative high-throughput screening (qHTS) initiative began in 2006 with 2800 compounds in >100 qHTS at 14 concentrations (5 nmol/L to 92 μmol/L) tested at the NIH Chemical Genomics Center (NCGC), which has been renamed the National Center for Advancing Translational Sciences (NCATS). The first Tox21 Memorandum of Understanding (MOU) was signed in early 2008 by the EPA, NIEHS/NTP, and National Human Genome Research Institute. A revised 5-year MOU was released in July 2010 and signed by the existing partners, and the FDA was added to the interagency project. Tox21 phase II was launched shortly thereafter using a library of 10 000 chemicals (10K Library) screened by approximately 30 assays per year. Phase II of the ToxCast program, which includes 800 compounds across 550 assays, is ongoing.

The goals of Tox21 are as follows: to identify patterns of compound-induced biological response in order to characterize toxicity/disease pathways, facilitate cross-species extrapolation, and model low-dose extrapolation; to explore the use of existing HTS assays for prioritization of chemical compounds for in vivo toxicological testing; and to develop predictive models for biological response in humans. Through application of these assays to a test set of chemicals with well-characterized toxicity profiles, the aim is to identify "predictive toxicity signatures or pathways" that can then be used to screen untested chemicals for potential biological activity of concern and prioritization for further, in vivo, testing.[13] In phase I of Tox21 and ToxCast, many of the assays used were proprietary assays designed to screen for potent compounds that exhibited targeted biological activity, similar to assays used in the pharmaceutical industry. The pharmaceutical approach necessitated that only 1 concentration of a commercially available compound be tested. By focusing on pharmaceutically active compounds, highly specific but relatively insensitive assays were selected. The challenge with these assays is their lack of applicability to chemical toxicity screening in which dose response of a pure and stabile compound is critical. Assay selection for Tox21 phase II is based on the following: information from in vivo toxicological investigations; phase I experience, advice of basic researchers, and

nominated assays; and maps of disease-associated cellular pathways. A critical issue for Tox21 is the need to identify human disease-associated pathways and useful assays for those pathways. Another critical issue for Tox21 is the need to extrapolate from acute to chronic exposure conditions and from in vitro concentration to in vivo dose.

Current State of the Science on HTS Assays in Hazard Identification

High-throughput screening data currently provide information on the biological activity of chemicals that may aid in prioritization and screening of chemical substances for hazards. The HTS data collection is in its initial stages as is the identification and selection of appropriate screening assays. As technologies advance and the screening assays are developed, they will be incorporated into the Tox21 effort (M. DeVito, personal communication, 2013).

The qHTS program is currently evaluating approximately 10 000 chemicals across a wide variety of biochemical assays. If one were to assume that 5% of those chemicals demonstrate biological activity, this would leave approximately 500 chemical substances requiring further evaluation in order to prioritize a more manageable number of chemicals. This scenario has called for the development of streamlined prioritization approaches. Streamlined approaches with mitochondrial disruption assays have indicated that a single HTS assay alone may not be sufficient for prioritization (M. DeVito, personal communication, 2013). Identification of false-positives and specific mechanisms of mitochondrial disruption in addition to the evaluation of tissue sensitivity using different cell types are other factors that will play a key role in prioritization. Prioritization should be based on risk not just hazard. Information on potential exposure levels in particular will be crucial in incorporating risk into prioritization efforts.

To better identify those chemicals that are of higher priority and to effectively develop targeted testing strategies for those chemicals, efforts are underway to evaluate the relationships (both qualitatively and quantitatively) between in vitro assays and in vivo biological activity and toxicity (M. DeVito, personal communication, 2013). Validation work focused on targeting testing in the liver, for example, is intended to test for the presence of in vivo activity as seen in in vitro assays and to confirm that previously untested compounds show predicted in vivo activity. Other approaches being researched to provide insight into the uncertainties in extrapolating HTS data to in vivo biological responses include in vitro to in vivo extrapolation via reverse toxicokinetics, short-term in vivo studies, orthogonal in vitro assays to help identify false-positives, and computational toxicology approaches (M. DeVito, personal communication, 2013).

Food-Relevant Compounds in HTS Assays

Of interest to the food industry is the extent to which food-relevant compounds are being assessed by HTS assays. To address this question, the ILSI North America Technical Committee on Food and Chemical Safety supported a summer fellowship program in which a doctoral student worked with the various regulatory agencies to identify the food-relevant chemicals in scope under Tox21 and to understand how the resulting data could be analyzed and interpreted (A. Forgacs, personal communication, 2013). Through two material transfer agreements (MTA) between the EPA and ILSI North America and the NIEHS and ILSI North America, the summer fellow was able to undertake the analysis of Tox21 data relevant to the food sector. Under the MTA with ILSI North America, the doctoral student was the only individual who had access to the Tox21 data.

To determine the extent to which food-relevant compounds are included in the current phase of the Tox21 program, HTS chemical inventories were mined for food-relevant chemicals including food additives, Generally Recognized as Safe (GRAS) substances, and food contact substances (FCS). The FDA publicly available databases were used as sources for the list of indirect and direct FCS, including the Everything Added to Food in the United States (EAFUS) list and the GRAS notification list. These lists were used as a reference to identify those chemical substances in the HTS database that are relevant to food. Over 5000 food-relevant chemicals were identified in which nearly 2000 were included in the qHTS phase II of the Tox21 10K Library and nearly 1000 in phase II of ToxCast.

In collaboration with a postdoctoral fellow at NIEHS/NTP, chemical and biological analysis involving the grouping of chemicals based on the physical/structural properties using Leadscope chemoinformatics software was conducted. This grouping provides a way to identify chemical classes that have similar biological activity. Once the chemical clusters were established, a modified Fisher's test was used to determine which assays were statistically relevant. This approach allows for the identification of assays that exhibit significant activity for food-relevant chemicals while identifying those structural clusters that elicit the greatest/least biochemical activity across the assays. It also provides valuable insight as to how such complex data can be managed and analyzed which may be of interest to a broad audience beyond the food industry.

In summary, HTS phase II of Tox21 currently includes many food-relevant chemicals. Some of those food-relevant chemicals have demonstrated biological activity in certain assays. Future work will seek to provide context for these findings by further evaluating the food-relevant compounds with respect to additional validation assays, published safety data, and biomonitoring datasets. This work will be discussed in more detail in the following section and is critical to assessing the significance of HTS data generated on food-relevant compounds. There is an opportunity to understand how HTS assays respond to food constituents, including beneficial nutrients, which by definition are biologically active substances. These compounds may play a critical role in interpreting and anchoring "toxicity signatures or pathways" that could be generated

by a variety of chemical substances by HTS assays (J. Bus and A. Forgacs, personal communication, 2013).

Enhancing the Confidence in HTS Assays for Hazard Identification

It is critical that the scientific community consider the extensive reservoir of knowledge gained from animal testing that would be helpful in framing future toxicological testing strategies (J. Bus, personal communication, 2013). Effective implementation of the Tox21 HTS testing strategies calls for recognition and understanding of dose response, organ-specific responses, multiorgan interactions, and mechanisms of action that drive the phenotypic expression of toxicity (J. Bus, personal communication, 2013). A significant concern with the reliance on in vitro assays is the identification of biological activity that is (1) disconnected from normal homeostatic or compensatory mechanisms present in vivo and (2) provoked due to high concentrations tested in vitro.[13] This concern reinforces the need for extensive validation of HTS assays in order to gain confidence that the HTS output is relevant to predicting hazards that may occur in vivo.

Based on the knowledge shared during the workshop, the HTS approach to prioritizing chemicals for toxicity testing is in its initial stages of scientific verification and development. Only through extensive rounds of HTS data collection, development of prioritization approaches and targeted testing will the scientific community be in a position to fully evaluate the utility of this approach (M. DeVito, personal communication, 2013).

Linking High-Throughput Screening Assays to Dosimetry

Speaker

Defining the Elements of a Holistic Weight of Evidence Approach.
James Bus, PhD, DABT, ATS, The Dow Chemical Company, Midland, MI, USA

High-throughput screening often involves the evaluation of chemicals at multiple test concentrations without clear consideration on how these in vitro test concentrations compare to doses used in animal studies and real-world human exposures. Advances in analytical and modeling technologies in both animal and human exposure evaluations (ie, biomonitoring) provide an opportunity to link in vitro test methods to whole animal toxicity through dosimetry (J. Bus, personal communication, 2013). By translating in vitro test concentrations to animal toxicology responses through "dosimetric anchoring," hazards irrelevant to human risk and nonreflective of whole animal metabolism can be identified.

The US EPA Endocrine Disruptor Screening Program (EDSP) uses default top-dose concentrations of 100 μmol/L or 1 mmol/L depending on the test system. The relationship between such in vitro default concentrations to in vivo animal dose and toxicity was explored with 2,4-dichlorophenoxyacetic acid (2,4-D). The no observable adverse effect level for 2,4-D based on animal studies is 0.7 μg/mL in plasma, while human biomonitoring data indicate that the real-world exposure to 2,4-D is 60 ng/mL in plasma. The concentration of 2,4-D tested in the EDSP in vitro screening assay was significantly higher (0.1 mmol/L equivalent to an estimated 22.1 μg/mL in human plasma) than the real-world exposure level. Although in vitro culture media and associated conditions are not directly parallel to blood serum or plasma (M. DeVito, personal communication, 2013),[14] these data show that the top tested concentration of 2,4-D in vitro was not in the range of physiologically relevant concentrations and thus was of limited if any value to informing the potential for human health risks.

A significant limitation to HTS screening data is the disconnect between in vitro test concentrations and exposure. By applying dosimetric anchoring to HTS screening data, false-positives and false-negatives for chemical hazard assessment can be avoided. While false-positives lead to unnecessary bans on valued compounds and misdirected resources away from issues of greater concern, false-negatives delay action on chemicals that are of concern. If the goal of HTS is to identify adverse outcomes of real-world concern, then these assays must be grounded by the exposure-to-dose paradigm.

Perspectives on Integrating Computational Modeling and HTS Assays into Risk Assessment

Speakers

Risk Assessment Roadmaps and Methods for Using 21st Century Methods.
Michael L. Dourson, PhD, DABT, ATS, Toxicology Excellence for Risk Assessment, Cincinnati, OH, USA

Perspectives from the Health Environmental Science Institute (HESI) Risk 21 Program.
Alan Boobis, OBE, PhD, CBiol, FSB, FBTS, Imperial College London, UK

Key Events Dose Response Framework (KEDRF).
Samuel Cohen, MD, PhD, University of Nebraska Medical Center, Omaha, NE, USA

Productive integration of increasingly complex data sets generated by in silico and HTS technologies into chemical risk assessments that improve the confidence of health evaluations will be challenging (A. Boobis, personal communication, 2013). To help address this challenge, a multisector, international team was assembled by the ILSI Health and Environmental Sciences Institute (HESI) to form the HESI Risk 21 initiative. The goal of HESI Risk 21 is to develop a scientifically based framework or roadmap that enables the incorporation of advances in 1.) computational and molecular toxicology, 2.) conventional toxicity testing and 3.) exposure in order to provide the

basis of an integrated evaluation strategy to chemical risk assessment (A. Boobis, personal communication, 2013).

The principles of HESI Risk 21 have been designed to bring applicable, accurate, and resource-appropriate approaches to the rapidly evolving world of human health risk assessment. This is best articulated by a shift in risk assessment philosophy away from conducting comprehensive toxicity testing to considering the problem that needs to be addressed and then selecting sources of information that will have the most value (A. Boobis and M. L. Dourson, personal communication, 2013). Wherever possible, prior knowledge on the physicochemical characteristics of the chemical, its exposure profile, and its toxicity as well as that of related compounds is used to determine the data needs for a new chemical that would be most informative (A. Boobis and M. L. Dourson, personal communication, 2013). Rather than assessing the hazard data first, the HESI Risk 21 program emphasizes the use of exposure estimates to provide insight into the toxicity data that may or may not be necessary.

In order to achieve this, HESI Risk 21 utilizes a tiered approach to obtain ranges for estimates of exposure and of toxicity in which the accuracy of those estimates increases with higher tiers. Estimates of exposure, for example, begin with tier 0 or worst-case approximations and then proceed to higher tiers with deterministic, probabilistic, and biomonitoring data. In terms of toxicity, assessments begin at the lowest tier such as QSAR and continue to progress to more significant data sets to in vitro, in vivo, and ultimately mode of action (MOA) data (A. Boobis, personal communication, 2013). The importance of MOA data highlights the critical role of the underlying biology in the expression of toxicity (M. L. Dourson, personal communication, 2013). The Key Events Dose Response Framework (KEDRF) is founded on MOA analysis, which is defined as a sequence of key events leading to the ultimate biological effect in an organism. The KEDRF focuses specifically on the dose response of each key event as a way to evaluate the ultimate dose response for the biologic effect being evaluated (S. Cohen and M. L. Dourson, personal communication, 2013).

The KEDRF is founded on first establishing a MOA by utilizing animal models. The Bradford Hill criteria are then applied to assess the strength of the MOA data including an analysis of dose–response relationships, temporal relationship, biological plausibility, and the strength, consistency, and specificity of the association of the toxicological effect with the key events identified (S. Cohen, personal communication, 2013). If the animal MOA meets these criteria, the key events within the MOA are then evaluated both qualitatively and quantitatively as to their relevance to humans. If any one of the key events leading up to the development of the ultimate biologic response was considered not to be relevant to humans, the MOA was also considered not to be relevant in humans (S. Cohen, personal communication, 2013). This level of analysis is consistent with the knowledge that each biologic effect will be evaluated individually even if multiple effects were due to the same MOA.

This approach has laid the foundation for a scientifically based framework in which inputs from in silico, HTS, and conventional toxicity data can form the basis of an integrated evaluation strategy for chemical risk assessment. It is important to note that while HTS and in silico data may certainly provide insight into MOA, the data need to be anchored to key events in animal studies in order to begin to assess whether the MOA can be extrapolated to humans (S. Cohen and M. L. Dourson, personal communication, 2013). The KEDRF framework also encourages transparency, provides guidance for data presentation, identifies critical data deficiencies and needs, and promotes global harmonization (S. Cohen, personal communication, 2013).

Recommendations for Optimizing Predictive Methods for Risk Assessment and Regulatory Use

Speaker

FDA Perspectives on the Application of 21st Century Toxicology Methods Used in Risk Assessment and Regulatory Decision Making.
David Hattan, PhD, FDA CFSAN, College Park, MD, USA

The risk assessment process outlined by the FDA begins with assessing the toxicological hazards of a compound by qualitative identification of possible adverse health effects (ie, hazard identification) followed by a quantitative assessment to determine the dose at which toxicity may occur (ie, hazard characterization). To put these findings into context, the likely exposure or intake is then estimated. This process then culminates in risk characterization in which there is a qualitative and quantitative estimation of the probability of occurrence and severity of known or potential adverse health effects in a given population based on the hazard identification, hazard characterization, and exposure assessment (Hattan, personal communication, 2013).

In conducting a review and determination of the safety of a food additive, the FDA weighs all of the available and relevant scientific evidence. The safety determination is not based on individual studies or isolated evidence. The mere presence of a compound or its metabolite in sample tissues or fluids is not determinative in and of itself of a health risk. Important criteria for assuring the validity and integrity of data used to support the safety of a chemical substance have been outlined in the FDA's Code of Federal Regulations (21 CFR) 58.185 (Hattan, personal communication, 2013). Of significant importance is a clear and detailed description of the methodology of a given technique, the statistical analyses, circumstances, or confounders that may have affected the quality or integrity of the data, all tabulated raw data, and an evidence-based interpretation of the data not to exceed the limits of the findings (Hattan, personal communication, 2013). One issue faced by many scientists who seek to validate HTS assays is the lack of visibility to HTS protocols used in some of the assays that are proprietary and cannot be shared. Without access to these

protocols, it is challenging to show that HTS assays meet the standards of data reliability and reproducibility as described by the FDA. Not only do the testing protocols themselves need to be validated but tests should also be conducted to determine whether the output correlates with animal studies (possibly via parallel study designs) to satisfy a representation of what occurs in vivo (Hattan, personal communication, 2013).

Panel Session: Practicality of Weight of Evidence: Where Do We Go From Here?

Moderator

Richard Canady, PhD, ILSI Research Foundation, Washington, DC, USA

Panelists

Roxi Beck, Center for Food Integrity
James Bus, PhD, The Dow Chemical Company, Midland, MI, USA
Keith Houck, PhD, EPA NCCT
Mike DeVito, PhD, NIH NIEHS NTP
James Rathman, PhD, Ohio State University, Columbus, OH, USA
Ken Wallace, PhD, University of Minnesota Medical School, Minneapolis, MN, USA

The workshop culminated in a panel discussion that revisited the following questions:

1. How would in silico and HTS data help in confirming the safety of approved food ingredients?
2. How do we begin the process to consistently and effectively use in silico and HTS data to help assess the safety of new food ingredient?
3. How would in silico and HTS data help in communicating the safety of food ingredients while enhancing the public's trust in the food supply?

In assessing these questions, the panelists agreed that the immediate goal was to use these methods to help prioritize chemicals for future testing, especially for those chemicals that have little data. It was highlighted that the real value from these methods could be derived by anchoring the output with animal and human data on exposure and health effects from approaches widely accepted by risk assessment and risk management experts. For example, if a compound has an in silico and HTS alert at estimated exposure levels relevant to a concordant concern identified in animal testing, then all lines of evidence could better inform future investigations. If concordance is not seen across the outputs from in silico, HTS, and animal testing, then additional work on validation and the need for more data are warranted.

It was clear that, as has always been the case, there are limits to the conclusions that can be based on the results of individual assays. The HTS and in silico modeling are currently experimental approaches to understanding biology, and no evidence was presented that supports the use of individual assay results as a basis for safety decisions. In fact, most participants agreed that there remains no clear consensus on the relevance of the output generated by these methods nor how the output should be applied in risk assessment. It has been suggested that in silico modeling can provide guidance in identifying potential chemical hazards, and validation efforts are being extended to new end points, and panelists generally supported this kind of approach. However, the use of HTS assays in identifying chemical hazards is still in its initial phases of development.

Two high-level needs were often repeated by the participants before input data could gain a position of proven utility in the risk assessment process. The first is the need to identify the relationship between in silico and HTS data to the possible exposure ranges of target chemicals through foods and the in vivo effects considering ADME. The second is the call for simple validation of the methods. Validation would include showing that, for example, ''off-the-shelf'' cell line assays developed for one purpose (drug discovery) are relevant to toxicity. It would also include measures of reproducibility with respect to the outcomes of interest (eg, predicting toxicity across a range of chemicals known to produce or not produce a particular effect). These are rough paraphrases of goals of the ToxCast and Tox21 programs, and it is hoped that over time they can be achieved. As mentioned previously, showing the relationship to safety assessment and demonstrating validity of these approaches begins with disclosure to the scientific community on the methods themselves in order to ensure reproducibility. All panelists expressed interest in advancing the scientific verification of these methods so that they could support safety assessments. Meeting this challenge will require an extensive and ongoing dialog between the toxicology, risk assessment, and other related scientific communities.

Communication

It is likely that the public release of data indicating alerts of chemical substances (including food-relevant substances) will be interpreted as risks to human health by the public and some organizations. Given the experimental nature of the methods, it is very likely that many if not most findings will have limited predictive value for actual health effects and will rather be informative of testing needs. However, at some point in the development of the methods, it seems likely that ''true'' predictors of harm may emerge. For this reason, the discussion and expectation of prediction of harm will continue for these methods. Considering this developmental nature of the meaning of the data, communication about the data requires careful consideration. Neither undue harm nor undue delay should be conveyed in such communication, as each may cause unwarranted effects to health through consumer response. For these reasons, consideration needs to be given to the key messages that accompany the data. In addition, it is important that transparency be considered by the owners of data and later users of the released data in their communications with the media and

the public. The Center for Food Integrity spoke about the need for communication that emphasizes 3 key aspects based on the shared values, trust, and science in order to build trust in today's food system in an era of radical transparency with the emergence of social media. According to the Center for Food Integrity, confidence or shared values are 3 to 5 times more important to consumers than scientific verification or competence in determining who they will trust in the food system. These findings should serve as a call-to-action for those in the government and the food system, and no longer it is sufficient to rely solely on science or to attack the attackers as a means of protecting self-interest. This new environment requires new ways of engaging and new methods of communication if the Tox21 community and industry want to build trust, earn and maintain social license, and protect the freedom to operate whether in a research laboratory or in business. Related to the theme of the workshop, the Center for Food Integrity said that we need stakeholders who control social license to understand that while our use of technology has increased and our food systems have changed, our commitment to doing what is right has never been stronger. The workshop participants were especially interested in using these communication concepts and approaches to the situation at hand where risk assessment methods are evolving in the absence of knowledge about their possible applications. The concern expressed was that the lack of understanding about these new input systems and the meaning of the data and results may be unfamiliar to many, and this knowledge void could be filled by opponents of the technology with misinformation, leading to a situation that would be very difficult to manage or reverse.

Based on this, it was recommended that industry and government stakeholders develop a communications strategy that focuses on the data to be released, format of the data to be released, and the shared commitment to the safety of the food supply. This strategy could begin with communicating how screening methods may contribute to food safety and the additional work that is underway to build confidence in these methods. For example, stakeholders could generate talking points, which state that these screening approaches are as follows: (1) a part of a scientific process to find new methods to analyze the safety of chemical substances, (2) represent a potential opportunity to further reduce animal testing, (3) remain under development and will require additional work to build confidence in their utility, and (4) signals of true health risk will be given high priority and addressed immediately. Simple messages are also needed to communicate that these assays are used for screening and cannot be used independently to determine safety and that despite what the name "Tox21" implies, the assays provide preliminary information on biological activity and not direct information on toxicity. The session of this workshop ultimately culminated in the following talking points:

- Tox21 is the beginning of an exciting, scientific process that could redefine the approach to assessing the safety of chemical substances.
- Despite what the name "Tox21" implies, the assays provide preliminary information on biological activity, not toxicity.
- Tox21 may be an opportunity to reduce animal testing but more work is required to confirm that the methods are accurate.
- Publishing the data and the results should include careful consideration to what the findings mean within the context of a comprehensive food safety assessment.
- The contribution of HTS data to a perception of public health risk will vary among various stakeholders.
- Balanced messaging concerning the results and the state of the science is unlikely to occur.
- In light of the release of HTS data, developing a communications strategy that the public will appreciate and understand is imperative.

Authors' Note

International Life Sciences Institute (ILSI) North America is a public, nonprofit foundation that provides a forum to advance understanding of scientific issues related to the nutritional quality and safety of the food supply by sponsoring research programs, educational seminars and workshops, and publications. ILSI North America receives support primarily from its industry membership. The views expressed in this research report are those of the authors and do not necessarily reflect the position or policy of PepsiCo, Inc.

Declaration of Conflicting Interests

The author(s) declared the following potential conflicts of interest with respect to the research, authorship, and/or publication of this article: HB is an employee of PepsiCo Inc.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship and/or publication of this article: Coauthor Jeffrey Barach supported the development of this article and received funds from ILSI North America for his work on this article. The workshop was sponsored by the North American Branch of the International Life Sciences Institute Technical Committee on Food and Chemistry Safety.

References

1. NRC. *Toxicity Testing in the 21st Century: A Vision and a Strategy*. Washington, DC: National Academies Press; 2007.
2. Memorandum of Understanding on High Throughput Screening, Toxicity Pathway Profiling, and Biological Interpretation of Findings between the US Department of Health and Human Services National Institutes of Health National Institute of Environmental Health Sciences/National Toxicology Program and the US Department of Health and Human Services National Institutes of Health National Human Genome Research Institute NIH Chemical Genomics Center, and the US Environmental Protection Agency Office of Research and Development; 2008. http://www.niehs.nih.gov/about/highlights/assets/docs/memorandum_of_understanding.pdf. Accessed July 25, 2013.

3. Patlewicz G, Simon T, Goyak K, et al. Use and validation of HT/HC assays to support 21st century toxicity evaluations. *Regul Toxicol Pharmacol*. 2013;65(2):259-268.
4. Hartung T. A toxicology for the 21st century—mapping the road ahead. *Toxicol Sci*. 2009;109(1):18-23.
5. Hartung T, van Vliet E, Jaworska J, Bonilla L, Skinner N, Thomas R. Systems toxicology. *ALTEX*. 2012;29(2):119-128.
6. Hartung T, McBride M. Food for Thought … on mapping the human toxome. *ALTEX*. 2011;28(2):83-93.
7. Hartung T, Luechtefeld T, Maertens A, Kleensang A. Integrated testing strategies for safety assessments. *ALTEX*. 2013;30(1):3-18.
8. Bidlack WR, Birt D, Borzelleca J, Clemens R, Coughlin JR. Expert report: making decisions about the risks of chemicals in foods with limited scientific information. *Comprehen Rev Food Sci Food Safety*. 2009;8(3):269-303.
9. Ferson S, Kreinovich V, Ginzburg L, Myers D, Sentz K. *Constructing Probability Boxes and Dempster-Shafer Structures*. Report SAND2002-4015. Albuquerque, NM: Sandia National Laboratories; 2003.
10. Sentz K, Ferson S. *Combination of Evidence in Dempster-Shafter Theory*. Report SAND2002-0835. Albuquerque, NM: Sandia National Laboratories; 2002.
11. Yager A, Liu L. *Classic Works on the Dempster-Shafer Theory of Belief Functions*. Berlin, Germany: Springer; 2007.
12. Dix DJ, Houck KA, Martin MT, Richard AM, Setzer RW, Kavlock RJ. The ToxCast program for prioritizing toxicity testing of environmental chemicals. *Toxicol Sci*. 2007;95(1):5-12.
13. Aylward LL, Hays SM. Consideration of dosimetry in evaluation of ToxCast data. *J Appl Toxicol*. 2011;31(8):741-751.
14. Blaauboer BJ. Biokinetic modeling and in vitro–in vivo extrapolations. *J Toxicol Environ Health B Crit Rev*. 2010;13(2-4):242-252.