Risk Analysis, Vol. 29, No. 4, 2009 DOI: 10.1111/j.1539-6924.2009.01222.x

From the Editors

Toxicity Testing in the 21st Century

This month's lead is a summary and review of the National Research Council's "Toxicity Testing in the 21st Century: A Vision and A Strategy."[1] In this article, Daniel Krewski, Melvin Andersen, Ellen Mantus, and Lauren Zeise summarize a vision to advance toxicity testing and human health assessment of environmental agents. They describe how scientific advances can transform toxicity testing to allow additional assessments of potentially toxic chemicals by using more timely and more cost-effective methods, including high- and medium-throughput *in vitro* screening assays, computational toxicology and systems biology, along with other emerging high-content testing methodologies, such as functional genomics and transcriptomics.

Suresh Moolgavkar, our Area Editor for Health Risk Assessment, asked six experts with different perspectives to comment on the paper. Each praises the vision and offers suggestions for making it more useful. Rory Connolly argues that if we expect risk assessment to maintain high throughput and be accurate, then there is need to address the issues of microdosimetry, adaptive responses and homeostasis. E. Donald Elliott focuses on the regulatory perspective, wondering why a regulator would ever take the political and legal risk to be the first to base an actual regulatory decision on the new model, and then he wonders if a judge would uphold a regulatory decision based on the new vision. Elliott argues for a legally sophisticated group or institution to take up the issues where the NAS Committee left off and fill in the gaps so that *Toxicity Testing in the 21st Century* can actually be used by regulatory authorities. Dale Hattis notes that while high-throughput testing may ultimately be of substantial value, for higher profile decisions on major agents in commerce involving complex tradeoffs of risks and economic impacts for different policy options, the findings of high-throughput tests will not be sufficient. For these, a system that quantitatively assesses actual health risks and the large associated uncertainties will be essential.

A commentary by Robert H. Kavlock and colleagues acknowledges the challenges laid out in the Krewski *et al.* perspective, and describes the NIH/EPA collaboration called *Tox21*. With four focus groups devoted to different components of the NRC vision, the *Tox21* consortium constitutes a concerted, long-term effort to identify mechanisms of chemically-induced biological activity, prioritize chemicals for more extensive evaluation and develop more predictive models of *in vivo* biological response. Lorenz R. Rhomberg urges *Risk Analysis* readers to read the full NRC report, focusing on a careful consideration of the ways that risk assessment will have to change to deal with the new testing approaches. He highlights his view that the new vision consists of more than new testing technologies, but is based on a change in the questions that toxicology addresses, that is, a shift toward identifying causes and then inferring possible effects. The final commentary by Joyce Tsuji stresses the difficulties of developing *in vitro* assays that can predict *in vivo* outcomes with adequate sensitivity and specificity and discusses challenges for public health decision-makers in dealing with uncertainty.

Krewski *et al.* briefly reply to each commentary and encourage us to use their paper and the accompanying commentaries as a starting point for thinking about a more complete evaluation of the future directions for toxicity testing as set out in the full NRC report. We're pleased with this set of papers and hope that you will consider proposing similar sets of papers to us.

The other papers in this issue examine terrorism, food contamination, endangered species, and other risk-related challenges. Yacov Haimes, our Area Editor for Engineering, had discussed the meaning of "vulnerability" in a 2006 article in *Risk Analysis.*[2] His perspective article in this issue examines what we mean by "resilience." He considers existing definitions and arrives at one that will prove useful to practitioners. Terje Aven and Ortwin Renn, funded by Norway's Research Council, consider the utility of

traditional quantitative risk assessments for complex risk issues, such as terrorism, in which the circumstances are both ambiguous and uncertain. They conclude that traditional Bayesian and other approaches provide an incomplete picture of risk. Instead, they suggest approaches grounded in qualitative methods that characterize the breadth of uncertainty and in scenarios validated through consistency, psychological empathy with the main players, congruence with past trends, and plausibility. Their scenario-generating approach includes processes aimed at enhancing imagination, using game theoretical experiments for simulating interactive variables and applying role playing for stimulating empathy. While favoring a more qualitative approach, they acknowledge that quantitative methods have a legitimate role to play in broader risk characterization and scenario approaches, as long as the limits with respect to their validity and reliability are noted.

Fear of food contamination is common, and we present four papers that address these increasingly high-profile risks. Funded by ZonMW and the Dutch Ministry of Public Health, Esther van Asselt *et al.* focus on cross-contamination and undercooking. They observed home preparation of chicken-curry salad, finding a wide range of microbial contamination levels in the final salad, caused by various cross-contamination practices and widely varying heating times. Model predictions indicated that cooking times should be at least 8 minutes and cutting boards need to be changed after cutting raw chicken in order to obtain safe bacterial levels in the final salad. The model predicted around 75% of the variance in cross-contamination. The authors suggest that model accuracy can further be improved by including other cross-contamination routes besides hands, cutting boards, and knives, and that the results be used as a worst-case estimate for evaluating cross-contamination in the home.

The first case of bovine spongiform encephalopathy (BSE) was detected in the United Kingdom in1986. Corporate and international ramifications are still occurring. Funded by the Dutch Ministry of Agriculture, Nature, and Food Quality, Clazien J. de Vos and Lourens Heres note that the ban on use of meat and bone meal (MBM) in livestock feed reduced the spread of BSE. But now that the BSE epidemic is fading out, should there be a partial lifting of the MBM ban? Developing a simulation model that considers three infection pathways, they find that cross-contamination in the feed mill is the most risky pathway. Combining model results, they conclude that the risk of using some MBM is very low.

Ides Boone, *et al.* focus on the quality of parameters and the probability of health risks associated with eating contaminated food. They select 101 parameters from the life cycle beginning with primary production and including transport, holding, and slaughterhouse; next to post-processing, distribution, and storage; and meal preparation and consumption.

Xiao-Wei Lin *et al.* construct a risk model to predict the diffusion of foot-and-mouth disease (FMD) caused by passengers carrying meat products from cloven-hoofed animals across international boundaries. Employing recent data from an international airport in Taiwan, they build a mathematical model to simulate the probability distributions of disease prevalence, of FMD virus existing in the meat products after meat processing, and they estimate the survival of virus. Notably, they report variations in illegal transport by season and that the odds of interception by trained animals are higher than by customs agents.

The remaining papers in this issue consider a variety of important risk-related issues and tools. Endangered species protection is a risk management issue in North America, and the precautionary principle is often cited as an ethical basis for protecting species. Funded by the U.S. National Science Foundation and Fisheries and Oceans Canada, Robin Gregory and Graham Long use a structured decision-making (SDM) approach to bound the management problem, define objectives and performance measures, develop precautionary management alternatives, and evaluate the consequences of management actions. They highlight how strategy tables were used by a stakeholder committee charged with protection of endangered Cultus Lake salmon on the Canadian west coast.

Tony Cox examines chronic obstructive pulmonary disease (COPD), which is the fourth leading cause of death worldwide. COPD is associated with smoking but many who smoke do not develop the disease and the disease continues to develop in those who stop smoking. Funded by Philip Morris, Cox builds a risk model based on the assumption of protease-antiprotease imbalance in the lung, leading to ongoing proteolysis (digestion) of lung tissue by excess proteases.

Funded partly by NIOSH, Robert Noble, A. John Bailer and Robert Park point out that model averaging in risk assessments has become more

common. Their paper tests a protocol for selecting models that best fit the data. The authors focused on evaluating the impact of coal dust on lung function in a cohort of exposed miners. After eliminating nearly all the models and combinations of models as not fitting their criteria, they observe that remaining models yield benchmark concentrations that differ by a factor of 2 to 9 depending on the concentration metric and covariates. Their approach is a useful strategy for addressing model uncertainty.

Hospitals are no longer considered safe places by many people. Funded by the French National Research Agency, Mathieu Emily *et al.* develop a coefficient that quantifies the risk associated with hospital departments, permitting comparisons of similar departments. The adjustment coefficient characterizes the tail of the distribution of the total patient length of stay in a department before the first disease event occurs. They provide an approximation for the tail of the distribution and illustrate it by examining the risk associated with a standard patient safety indicator in 20 southeastern French hospitals.

Finally, Natarajan Krishnamurthy reviews *Risk: Philosophical Perspectives* (2007), edited by Tim Lewens, and recommends the collection of essays for those interested in the philosophical and psychological underpinnings of risk assessment and management.

Michael Greenberg
Karen Lowrie

REFERENCES

1. National Research Council. Toxicity Testing in the 21st Century: A Vision and a Strategy. Washington, DC: National Academy Press, 2007.
2. Haimes, YY. On the definition of vulnerabilities in measuring risks to infrastructures, Risk Analysis, 2006;26(2): 293–296.

Risk Analysis, Vol. 29, No. 4, 2009 DOI: 10.1111/j.1539-6924.2008.01150.x

Perspective

Toxicity Testing in the 21st Century: Implications for Human Health Risk Assessment

Daniel Krewski,[1]* Melvin E. Andersen,[2] Ellen Mantus,[3] and Lauren Zeise[4]

At the request of the Environmental Protection Agency, the National Research Council (NRC) recently completed a major report entitled *Toxicity Testing in the 21st Century: A Vision and a Strategy*. The terms of reference for this report were to develop a long-range vision and strategic plan to advance the practices of toxicity testing and human health assessment of environmental agents. The report describes how current and anticipated scientific advances can be expected to transform toxicity testing to permit broader coverage of the universe of potentially toxic chemicals to which humans may be exposed, using more timely and more cost-effective methods for toxicity testing. The report envisages greatly expanded use of high- and medium-throughput *in vitro* screening assays, computational toxicology, and systems biology, along with other emerging high-content testing methodologies, such as functional genomics and transcriptomics. When fully implemented, the vision will transform the ways toxicity testing and chemical risk assessment are conducted, moving away from measuring apical health endpoints in experimental animals toward identification of significant perturbations of toxicity pathways using *in vitro* tests in human cells and cell lines. Population-based studies incorporating relevant biomarkers will also be useful in identifying pathway perturbations directly in humans and in interpreting the results of *in vitro* tests in the context of human health risk assessment. The present article summarizes and extends the NRC report and examines its implications for risk assessment practice.

1. BACKGROUND

Toxicity testing has traditionally relied on studies of adverse health outcomes observed in animals at high doses, with subsequent extrapolations to expected human responses at much lower doses. These approaches date back to the 1950s, when knowledge of the biology underlying toxic response was primitive. A recent report from the National Academy of Sciences, *Toxicity Testing in the 21st Century*,[1] has proposed fundamentally new directions for toxicity testing in light of advances in understanding biological responses to chemical stressors. The vision was motivated by the need to greatly expand the coverage of the universe of environmental agents to which we are potentially exposed, ideally generating relevant toxicity data on most, if not all, such agents in a timely and cost-effective manner suitable for human health risk assessment, including vulnerable populations.

The vision rests on a fundamentally new approach to toxicity testing based on evaluation of perturbations of toxicity pathways identified using a comprehensive suite of high-throughput *in vitro*

[1] University of Ottawa, Ottawa, Canada.
[2] Hamner Institutes for Health Research, Research Triangle Park, NC, USA.
[3] National Research Council, Washington, USA.
[4] California Environmental Protection Agency, Oakland, USA.
* Address correspondence to D. Krewski, Institute of Population Health, University of Ottawa, Room 320, 1 Stewart St., Ottawa, KIN 6N5, Canada; cphra@uottawa.ca.

assays in human cells and cell lines. Dose-response assessments would often involve computational models of the circuitry of these pathways and pharmacokinetic models to support *in vitro* to *in vivo* extrapolations. Risk assessment would focus on maintaining exposures to environmental agents below the level at which significant perturbations of these pathways could occur. Modern biology provides the tools needed to make this vision a reality.

2. FRAMING THE VISION

In developing its vision for toxicity testing, the committee considered four options for toxicity testing (Table I). In choosing among these options, the committee sought to define a toxicity-testing paradigm that would (1) achieve broad coverage of chemicals, chemical mixtures, outcomes, and life stages, (2) reduce the cost and time required for toxicity testing, (3) develop a more robust scientific basis for assessing health effects of environmental chemicals, and (4) minimize use of animals in testing.

Option I (essentially the status quo) retains current toxicity-testing principles and practices, relying primarily on *in vivo* animal toxicity tests to predict human health risks. Under this option, the difficulties in interpreting animal data obtained at high doses with respect to risks in the heterogeneous human population would remain. Reliance on whole animal testing is both expensive and time consuming, thereby limiting toxicity-testing throughput.

Tiered testing (Option II) can be used to generate pertinent data for a more efficient assessment of potential health risks of environmental agents, taking into consideration available knowledge on the chemical and its class, its mechanisms of action, and its intended use and estimated exposures.[2] These factors are used in refining testing priorities to focus first on areas of greatest concern in early tiers and then judiciously moving to advanced testing in subsequent tiers, as needed. This option represents a small step in improving coverage, reducing costs and animal usage, and increasing the use of mechanistic information in risk assessment.

A more transformative paradigm shift is needed to achieve the objectives for toxicity testing (Option IV). The committee's vision for accomplishing this goal is built upon the identification of biological perturbations in toxicity pathways that can lead to adverse health outcomes in humans under conditions of human exposure. The use of a comprehensive array of *in vitro* tests of toxicity pathway responses to identify relevant biological perturbations using cellular and molecular systems based on human biology could eventually eliminate the need for whole animal testing and provide much stronger, mechanistically based predictive tools for human health risk assessments. *In silico* computational screens could also play a role in predicting the likelihood that a particular environmental agent would cause adverse health outcomes in humans, although further testing would likely be pursued.

This new approach would be less expensive and less time consuming than the current approach, resulting in much higher throughput via use of modern robotic methodologies for pathway testing. Although the reliance on *in vitro* results for human health risk assessment lacks the whole organism integration provided by current tests, risk assessments would be based on avoiding biological perturbations in toxicity pathways that can reasonably be expected to lead to adverse health effects.

Clearly, there are a number of scientific challenges to fully implementing this vision. Major

Table I. Options for Future Toxicity-Testing Strategies

Option I *In Vivo*	Option II Tiered *In Vivo*	Option III *In Vitro*/*In Vivo*	Option IV *In Vitro*
Animal biology	Animal biology	Primarily human biology	Primarily human biology
High doses	High doses	Broad range of doses	Broad range of doses
Low throughput	Improved throughput	High and medium throughput	High throughput
Expensive	Less expensive	Less expensive	Less expensive
Time consuming	Less time consuming	Less time consuming	Less time consuming
Relative large number of animals	Fewer animals	Substantially fewer animals	Virtually no animals
Apical endpoints	Apical endpoints Some *in silico* and *in vitro* screens	Perturbations of toxicity pathways *In silico* screens possible	Perturbations of toxicity pathways *In silico* screens

Source: Adapted from NRC.[1] Reprinted with permission; copyright 2007, National Academies Press.

concerns relate to ensuring adequate testing of metabolites and the potential difficulties in evaluating novel chemicals, such as nanomaterials and biotechnology products, using existing *in vitro* tests. These challenges may require maintenance of some whole animal tests into the foreseeable future, as reflected in Option III. This option also leaves open the possibility of more extensive *in vivo* evaluations of the toxicity of representative chemicals from entirely new classes of environmental agents.

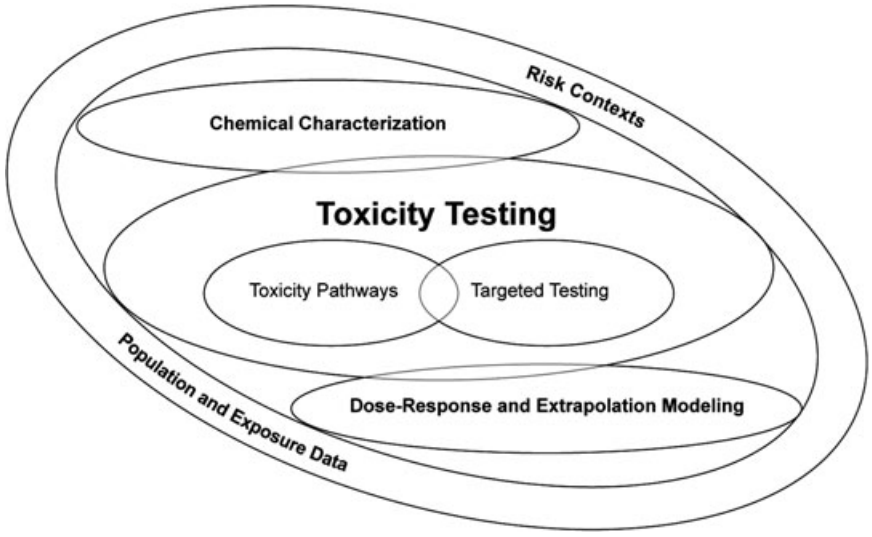
3. COMPONENTS OF THE VISION

The main components of the committee's vision are shown in Fig. 1. *Chemical characterization* involves the collection and computational modeling of physical and chemical properties, use, environmental concentrations and stability, likely routes of human exposure, the potential for bioaccumulation, metabolites and breakdown products, molecular interactions with cellular components, and potential toxic properties.

Rapid, inexpensive tests for perturbations of *toxicity pathways* are central to the vision for toxicity testing. The vision emphasizes the development of a broad suite of high-throughput toxicity pathway assays that use primary cells or cell lines, preferably of human origin, to evaluate relevant perturbations.[3,4] The efficiencies afforded by use of high-throughput *in vitro* screens will permit not only broader coverage of the universe of environmental agents that may pose a risk to human health, but also of mixtures of environmental agents. As *in vitro* tests become more refined, such assays may also facilitate investigation of toxic effects at different life stages and in vulnerable populations demonstrating genetic or other susceptibilities to exposure to an environmental agent.

Until prediction of metabolism can be more reliably accomplished through computational toxicology and *in vitro* testing, *targeted testing* using whole animals will likely be needed to identify toxic metabolites that require evaluation by high-throughput testing. Other in-life testing may still be required to clarify substantial uncertainties in the interpretation of toxicity pathway data or to fill gaps in the toxicity pathway testing strategy to ensure that critical toxicity pathways and endpoints are adequately covered.

Dose-response and extrapolation modeling should provide integrative tools for interpreting toxicity-testing data.[5] Dose-response analysis will be greatly informed by the use of computational systems biology to describe biological circuitry of toxicity pathways.[6] Pharmacokinetic models[7] will be needed to determine environmental exposure levels giving rise to human tissue concentrations comparable to those associated with perturbations of toxicity pathways *in vitro*.[8] Human data would provide information on background chemical



Source: NRC.[1] Reprinted with permission; copyright 2007, National Academies Press.

Fig. 1. The proposed vision for toxicity testing includes chemical characterization, toxicity testing, and dose-response and extrapolation modeling. At each step, population-based data and human exposure information are considered in the context of the data needed for decision making.

exposures and disease processes that would affect the toxicity pathway and provide a basis for addressing host susceptibility.

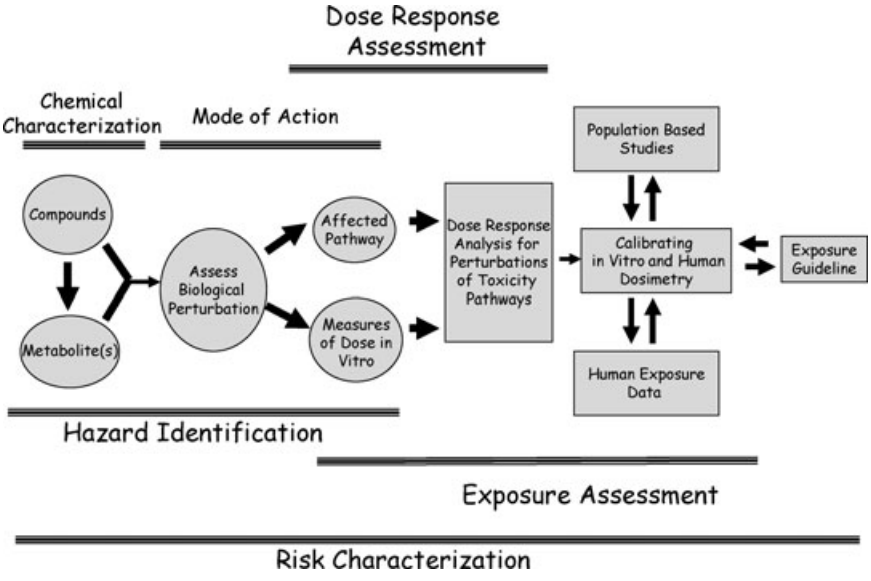
The vision includes the generation and use of *population-based and human exposure data* for interpreting toxicity-test results and encourages the collection of such data through human biomonitoring, environmental health surveillance, and targeted epidemiologic studies, particularly those involving molecular and genetic components.[9] *In vitro* toxicity tests conducted in human cells can help identify specific biomarkers of exposure, biologic change, or susceptibility that can be investigated directly in human populations. Our expanding knowledge of the biological mechanisms underlying human disease will be valuable in charting toxicity pathways of relevance for human health risk assessment.

Risk management decision making may vary according to the *risk context* that creates the need for toxicity-testing information. Commonly encountered scenarios include evaluation of potential environmental agents, existing environmental agents, sites of environmental contamination, environmental contributors to a human disease, and the relative risk of different environmental agents.

4. TOXICITY TESTING AND RISK ASSESSMENT

The ultimate application of the results of toxicity testing is in the risk assessment of environmental agents. The committee was charged with developing a vision for toxicity testing that would better inform the assessment of the potential human health risks of exposure to environmental agents for which the EPA is responsible, as well as ensuring the existence and availability of efficient test methods. The committee's vision is based on a shift away from traditional toxicity testing focusing on the demonstration of adverse health effects in experimental animals toward a deeper understanding of biological perturbations in key toxicity pathways leading to adverse health outcomes.

As illustrated in Fig. 2, there is a close linkage between the risk assessment paradigm originally put forward by the NRC[10] in its "red book" and the committee's vision for toxicity testing. Chemical characterization and mode of action analyses are involved in hazard characterization. Mode of action analysis involves the application of a wide range of methods, including computational chemistry and *in vitro* screening. Genomics, computational systems biology, and biologically based dose-response models offer powerful tools for describing dose-response relationships. Pharmacokinetic models can be used to calibrate *in vitro* and human dosimetry, thereby facilitating the translation of dose in cellular systems to dose in human organs and tissues. Population-based studies may be used to interpret findings from *in vitro* studies in the context of human health risk assessment and to identify critical toxicity pathway perturbations directly in humans using appropriate



Source: Adapted from NRC.[1] Reprinted with permission; copyright 2007, National Academies Press.
Fig. 2. Toxicity testing and risk assessment.

biomarkers. Collectively, this work permits the establishment of human exposure guidelines designed to avoid critical pathway perturbations.

5. CHALLENGES IN IMPLEMENTING THE VISION

The vision presented by the NRC will face a number of challenges in its implementation. The reliance on toxicity pathway perturbations as the basis for human health risk assessment will require sufficient understanding of such pathways to permit the shift away from apical outcomes in animals to occur with confidence. A major research effort conducted over a period of a decade or more is needed to effectively "map" all of the important toxicity pathways in humans, similar to the recently completed effort to map the human genome. As part of this research effort, a clear distinction will need to be made between unimportant biological changes and critical pathway perturbations that can, if not prevented, be expected to lead to adverse health outcomes.

Implementation of the vision will require a concerted effort to develop both the comprehensive suite of toxicity pathway assays and the regulatory and political infrastructure enabling their use in health risk assessment. There will be technical challenges in developing and refining the toxicity-testing tools and technologies that will provide the toxicity pathway assays on which the vision rests. However, recent advances in computational toxicology under the EPA's ToxCastTM program[11] and high-throughput screening at the National Institute for Environmental Health Sciences[12] and the National Chemical Genomics Center[13] provide a strong basis for optimism in this regard. The use of such high-throughput screens will facilitate the testing of larger numbers of doses spanning a broader range, including those relevant to human exposure conditions. As sensitive screens are developed, it will be possible to obtain more precise information on toxicity pathway perturbations at low doses than is currently the case for low-dose apical responses.

Successful implementation of the vision will also require that the tests for toxicity pathway perturbations be validated. Although both the Interagency Coordinating Committee for Validation of Alternative Methods (ICCVAM) and the European Center for the Validation of Alternative Methods (ECVAM) have well organized programs for the validation of nonanimal test systems, the focus of these programs is on validation against animal test results.[14] As animal tests are replaced by *in vitro* tests for perturbations of toxicity pathways, it will be important to validate the *in vitro* tests with respect to their relevance for humans.

Regulatory authorities will need to consider how current risk assessment practices can be adapted to make use of the types of toxicity-testing data underlying the committee's vision. Lawmakers will need to exercise flexibility in the interpretation of regulatory statutes, such as the Toxic Substances Control Act, or possibly update them, to reflect the reliance on biologically significant perturbations of key toxicity pathways, rather than adverse effects associated with maintenance of these perturbations over extended periods of time in experimental animal tissues.[5]

The committee concluded that an appropriate institutional structure for the proposed vision is an interdisciplinary research institute, coordinated and funded primarily by the federal government, that fosters intramural and extramural research. Given the scientific challenges and knowledge development required, significant funding will be needed to implement the vision over the next one to two decades;[6] the report envisions a research and test development program similar in scale to the National Toxicology Program. Without this investment, toxicity testing will evolve at a much slower pace, leaving the limitations of the current paradigm largely intact for decades to come. This investment will produce more relevant data on which assessments of human health risk can be based and greatly expand the coverage of the numbers of chemicals that can be tested. These improvements will strengthen our ability to protect people from the potential risks posed by environmental agents and permit continuing incorporation of new knowledge about toxicity pathways and their function into a modern toxicity-testing paradigm geared toward quantitative assessment of human health risks at relevant exposures.

[5] The need to update existing federal regulatory statues was debated in a session organized by the Environmental Law Institute, hosted by the District of Columbia Bar Association on December 11, 2007. Initial opinion expressed by legal experts was that the risk assessment implications of the vision could be accommodated through appropriate rulemaking, without the need to revise the statutes *per se*.[15]

[6] Since the time course required to fully implement the vision is difficult to predict with precision, review of progress, with appropriate mid-course corrections, was also recommended.[1]

6. CONCLUSION

The vision for toxicity testing in the 21st century articulated by the NRC[1] represents a paradigm shift away from adverse effects observed in experimental animals at high doses toward identifying and avoiding biologically significant perturbations of key toxicity pathways. The vision advocates new approaches to toxicity testing within each of its main components, which, collectively, will ultimately lead to a transformation in the way the potential health risks of environmental agents are assessed. The U.S. Environmental Protection Agency, the NIH Chemical Genomics Center, and the National Institute of Environmental Health Sciences recently announced a collaborative program to make the vision presented here a reality.[16] Further engagement of the broader scientific community, along with additional discussion and debate involving all stakeholder groups, will be invaluable in achieving this goal.

ACKNOWLEDGMENTS

This communication is based in large part on the final report of the NRC Committee on Toxicity Testing and Assessment of Environmental Agents.[1] The committee was comprised of the following members: Daniel Krewski (*Chair*), Daniel Acosta, Jr., Melvin Andersen, Henry Anderson, John Bailar III, Kim Boekelheide, Robert Brent, Gail Charnley, Vivian Cheung, Sidney Green, Karl Kelsey, Nancy Kerkvliet, Abby Li, Lawrence McCray, Otto Meyer, D. Reid Patterson, William Pennie, Robert Scala, Gina Solomon, Martin Stephens, James Yager, Jr., and Lauren Zeise. The NRC project director was Ellen Mantus. D. Krewski is the NSERC/SSHRC/McLaughlin Chair in Population Health Risk Assessment at the University of Ottawa. The authors are grateful to two referees for their helpful comments on the original draft of this article.

REFERENCES

1. National Research Council. Toxicity Testing in the 21st Century: A Vision and a Strategy. Washington, DC: National Academy Press, 2007.
2. Carmichael NG, Barton HA, Boobis AR, Cooper RL, Dellarco VL, Doerrer NG, Fenner-Crisp PA, Doe JE, Lamb JC, Pastoor TP. Agricultural chemical safety assessment: A multisector approach to the modernization of human safety requirements. Critical Reviews in Toxicology, 2006; 36:1-7.
3. Inglese J, Auld DS, Jadhav A, Johnson RL, Simeonov A, Yasgar A, Zheng W, Austin CP. Quantitative high-throughput screening: A titration-based approach that efficiently identifies biological activities in large chemical libraries. Proceedings of the National Academy of Sciences of the United States of America, 2006; 103:11473-11478.
4. Austin CP, Brady LS, Insel TR, Collins FS. NIH molecular libraries initiative. Science, 2004; 306:1138-1139.
5. Andersen ME, Yang RS, French CT, Chubb LS, Dennison JE. Molecular circuits, biological switches, and nonlinear dose-response relationships. Environmental Health Perspectives, 2002; 110(Suppl. 6):971-978.
6. Alon U. An Introduction to Systems Biology: Design Principles of Biological Circuits. Boca Raton, FL: CRC Press, 2006.
7. Reddy MB, Yang RSH, Clewell HJ, III, Andersen ME (eds). Physiologically Based Pharmacokinetics: Science and Application. Hoboken, NJ: John Wiley & Sons, 2005.
8. Blaauboer BJ. The integration of data on physico-chemical properties, *in vitro*-derived toxicity data and physiologically based kinetic and dynamic as modelling a tool in hazard and risk assessment, A commentary. Toxicology Letters, 2003; 138:161-171.
9. National Research Council. Human Biomonitoring for Environmental Chemicals. Washington, DC: National Academy Press, 2006.
10. National Research Council. Risk Assessment in the Federal Government: Managing the Process. Washington, DC: National Academy Press, 1983.
11. Dix DJ, Houck KA, Martin MT, Richard AM, Setzer RW, Kavlock RJ. The ToxCast program for prioritizing toxicity testing of environmental chemicals. Toxicological Sciences, 2007; 95:5-12.
12. Xia M, Huang R, Witt KL, Southall N, Fostel J, Cho MH, Jadhav A, Smith CS, Inglese J, Portier CJ, Tice RR, Austin CP. Compound cytotoxicity profiling using quantitative high-throughput screening. Environmental Health Perspectives, 2008; 116:284-291.
13. Titus SA, Xiao L, Southall N, Lu J, Inglese J, Brasch M, Austin CP, Zheng W. A cell-based PDE4 assay in 1536-well plate format for high-throughput screening. Journal of Bimolecular Screening 2008 [Epub ahead of print].
14. Kirkland D, Pfuhler S, Tweats D, Aardema M, Corvi R, Darroudi F, Elhajouji A, Glatt H, Hastwell P, Hayashi M, Kasper P, Kirchner S, Lynch A, Marzin D, Maurici D, Meunier JR, Müller L, Nohynek G, Parry J, Parry E, Thybaud V, Tice R, van Benthem J, Vanparys P, White P. How to reduce false positive results when undertaking *in vitro* genotoxicity testing and thus avoid unnecessary follow-up animal tests: Report of an ECVAM workshop. Mutation Research, 2007; 628:31–55.
15. Cohen BC, Crofton KM, Hartung T, Krewski D, Locke PA, Stephens ML. (individual commentaries). Toxicity testing in the 21st century: Better results, less use of animals. Environmental Forum, 2008; 25:46-51.
16. Collins FS, Gray GM, Bucher JR. Transforming environmental health protection. Science, 2008; 319:906-907.

Risk Analysis, Vol. 29, No. 4, 2009 DOI: 10.1111/j.1539-6924.2008.01165.x

Commentary

Commentary on "Toxicity Testing in the 21st Century: Implications for Human Health Risk Assessment" by Krewski *et al.*

Rory B. Conolly*

The main goal of *Toxicity Testing in the 21st Century: A Vision and a Strategy*[1] is to provide a practical means of evaluating the heath risks of chemicals. The vision is in stark contrast to the current situation, where the focus on apical testing *in vivo* is too expensive and too slow to keep up with ongoing influx of new chemicals into commerce. Key aspects of the vision are (a) *in vitro* evaluation of perturbations of toxicity pathways rather than *in vivo* evaluation of apical effects and (b) evaluation of dose response over relevant concentration ranges, as opposed to the current practice of focusing on doses that cause adverse effects in laboratory animals. While the NAS report lacks specifics on how the vision should be implemented, and even a precise definition of a toxicity pathway, its concern for a practical and dose-relevant approach to risk assessment is laudable.

The central concern of risk assessment is to ensure that the public health is not adversely affected by exposure to chemicals and other environmental agents. It follows that risks should be predicted as accurately as possible, since accurate prediction helps to ensure that (a) the associated exposure standards actually do protect the public health and (b) the economic consequences of compliance with the standards are commensurate with the risk. In today's world we have little assurance that predicted risks are accurate, and while it is possible to argue that current practice is health protective, it is unlikely that exposure standards and their economic consequences are commensurate with actual risk. Since remediation is expensive,[2] assurance that cost is aligned with risk is necessary if we are to avoid the situation where inaccurate risk assessment results in misallocation of economic resources.

Strikingly absent from Krewski *et al.*[3] and from *Toxicity Testing in the 21st Century: A Vision and a Strategy*[1] is any clear-cut statement of concern for the accuracy of predicted risks. This raises the possibility that the vision could be implemented in a manner that achieves high throughput, and that focuses on relevant doses, but that at the end of the day does not lead to more accurate risk assessment. Concern for protection of the public health, and for unnecessary economic consequences, would remain. This result is likely if the vision is implemented with a less-than-adequate understanding of the quantitative relationships between acute evaluations of *in vitro* perturbations of toxicity pathways and development of apical effects in humans, both acute and chronic. What factors should be considered to avoid such an unfortunate outcome?

The vision clearly articulates the importance of an understanding of exposure and of the relationship between exposure and tissue dose. However, the focus on *in vitro* tests suggests that estimates of dose at the level of individual cells will be needed, not the average dose to all the cells in a tissue. We know very well that concentration gradients across tissues can be steep, as shown by Kimbell *et al.*[4] for the respiratory tract, and that important dose-dependent transitions occur in the activation of toxicity pathways.[5] Similar observations have been made for other tissues, including the liver[6] and kidney.[7] Implementation of the vision will thus

National Center for Computational Toxicology, Office of Research and Development, U.S. Environmental Protection Agency, MD B205-01, Research Triangle Park, NC, USA.

*Address correspondence to Rory B. Conolly, National Center for Computational Toxicology, Office of Research and Development, U.S. Environmental Protection Agency, MD B205-01, Research Triangle Park, NC 27711, USA; tel: 919.541.3350; Conolly.Rory@epa.gov.

require significant refinement in the current state of the art for both pharmacokinetic data collection and for dosimetry modeling. Today, average tissue doses are typically measured and dosimetry models consist of homogeneous, well-stirred compartments. The vision, however, will require some combination of "microdosimetry" data at the level of the individual cells and a robust ability to computationally predict such data. In the future, pharmacokinetic models that incorporate spatial information on tissue architecture may provide this predictive capability at the cellular level or, at a minimum, at some greater level of resolution than is typically available today. It should be noted that this need is already starting to be addressed in projects whose long-range goals are the development of "virtual tissues," which are in silico descriptions of tissues across multiple levels of biological organization.[8]

Beyond the issue of dosimetry, it is important to ask how accurately, in the quantitative sense, we can expect *in vitro* assays to toxicity pathways to recapitulate apical toxicity *in vivo*. A central concern here is that the quantitative tissue dose-apical toxicity relationship *in vivo* reflects adaptive responses designed to maintain homeostasis[9,10] and that these adaptive responses may depend on higher levels of biological organization that will not be captured in *in vitro* assays. For example, the dose response for apical toxicities that are sensitive to steroid hormones will be affected by feedback signaling across the multiorgan endocrine system. It is difficult to envision an in *vitro test*, or even a suite of *in vitro* tests, that would adequately recapitulate such system-wide behavior. Here again, virtual tissues may prove be a key enabling technology, since they will in theory be able to integrate cellular-level data to higher levels of biological organization. Perhaps with time it will be possible to construct systems of linked virtual tissues capable of predicting dose response *in vivo* and of providing mechanistic insight into discrepancies between dose responses measured *in vitro* and *in vivo*. Such a capability is only a theoretical possibility today, but this is not inconsistent with the long-range nature of the vision and the recognition that substantial funding over a 10–20 year timeframe, or perhaps even longer, will be needed for its implementation.[1]

In summary, issues that will need to be considered if the NAS/NRC vision is to be successfully implemented have been discussed. If we are to strive for risk assessment that is both high throughput and accurate, the issues of microdosimetry, adaptive responses, and homeostasis will need to be addressed. It is hoped that this commentary will be viewed not as a negative criticism of the vision but, rather, as encouragement intended to help focus the vision along a pathway leading to its successful implementation.

DISCLAIMER

This work was reviewed by the U.S. EPA and approved for publication but does not necessarily reflect EPA policy.

REFERENCES

1. National Research Council. Toxicity Testing in the 21st Century: A Vision and a Strategy. Washington, DC: National Academy Press, 2007.
2. Fields S. The earth's open wounds. Environmental Health Perspective, 2003; 111:A154-A161.
3. Krewski D, Andersen ME, Mantus E, Zeise L. Toxicity testing in the 21st century: Implications for human health risk assessment. Risk Analysis, 2008 (in press).
4. Kimbell JS, Subramaniam RP, Gross EA, Schlosser PM, Morgan KT. Dosimetry modeling of inhaled formaldehyde: Comparisons of local flux predictions in the rat, monkey, and human nasal passages. Toxicological Science, 2001; 64:100-110.
5. Slikker W Jr, Andersen ME, Bogdanffy MS, Bus JS, Cohen SD, Conolly RB, David RM, Doerrer NG, Dorman DC, Gaylor DW, Hattis D, Rogers JM, Setzer RW, Swenberg JA, Wallace K. Dose-dependent transitions in mechanisms of toxicity. Toxicology and Applied Pharmacology, 2004; 201:203-225.
6. Jones AL, Hradek GT, Renston RH, Wong KY, Karlaganis G, Paumgartner G. Autoradiographic evidence for hepatic lobular concentration gradient of bile acid derivative. American Journal of Physiology. Gastrointestinal and Liver Physiology, 1980; 238:G233-G237.
7. Stern L, Backman KA, Hayslett JP. Effect of the cortical-medullary gradient for ammonia on urinary excretion of ammonia. Kidney International, 1985; 27:652-661.
8. Kavlock RJ, Ankley G, Blancato J, Breen M, Conolly R, Dix D, Houck K, Hubal E, Judson R, Rabinowitz J, Richard A, Setzer RW, Shah I, Villeneuve D, Weber E. Computational toxicology—A state of the science mini review. Toxicological Science, 2008; 103:14-27.
9. Li N, Nel AE. Role of the Nrf2-mediated signaling pathway as a negative regulator of inflammation: Implications for the impact of particulate pollutants on asthma. Antioxidants and Redox Signaling, 2006; 8:88-98.
10. Zhang Q, Andersen ME. Dose response relationship in antistress gene re24-egulatory networks. PLoS Computational Biology, 2007; 3:345-363.

Risk Analysis, Vol. 29, No. 4, 2009 DOI: 10.1111/j.1539-6924.2008.01166.x

Commentary

Needed: A Strategy for Implementing the Vision

E. Donald Elliott*

The title of the NAS Report was *Toxicity Testing in the 21st Century: A Vision and a Strategy*. Both the NAS report itself, and the current synopsis, are long on the "vision thing" but short on the "strategy" part. What is needed next is more thought about how to turn this provocative vision into a reality for actually implementing the new paradigm for toxicity testing in regulatory risk assessments upon which important health and safety decisions are based.

How can the new paradigm actually become accepted in regulatory decision making? For example, are we ready to approve new drugs for market, or set regulatory limits for environmental agents, based solely on computational models of toxicity pathways? If the answer is "yes" in some cases, but "no" in others, how are we to distinguish the two? If the answer is "no, not yet," what's missing and how will we know when we are ready?

This is not to minimize the accomplishments either of the NAS Committee, or of the present authors in summarizing this very important vision for the future of toxicity testing succinctly and accurately. The Committee also deserves credit for going beyond merely outlining a scientific vision. The title of the final chapter of the toxicity testing report promises to address "Prerequisites for Implementing the Vision in Regulatory Contexts," but the institutional and legal issues are treated in a perfunctory and superficial way. The full report, like the Perspective, largely contents itself with the observation that perhaps the new paradigm could be implemented under the language of existing statutes. While I agree with that conclusion, it is merely the beginning, not the end, of a serious discussion of regulatory implementation. How is a regulator to decide when to use the new paradigm as opposed to the prevailing whole body animal testing model? Why would a regulator ever take the political and legal risk to be the first to base an actual regulatory decision on the new model? Why would a judge ever uphold a regulatory decision based on the new vision? Should judges admit expert testimony based on the new paradigm in toxic tort cases under the *Daubert* rule that requires scientific evidence to be "reliable" and "valid"?[1]

Perhaps there are good answers to these and other implementation questions. Ambitious new testing programs such as REACH are literally impossible without new methods for screening and prioritizing. My colleague Bret Cohen has suggested that EPA could convene a rulemaking under §4 of the Toxic Substances Control Act to consider in what areas, if any, the new vision may be ready for actual use in practice, and if so, under what circumstances.[2] Perhaps we need a period of "ground truthing" in which both whole body animal testing and the new vision for toxicity pathways are applied side-by-side and the results compared. Perhaps we need to develop verifiable criteria for when the new vision should and should not be used.

The missing link needed now is for a legally sophisticated group or institution to take up the issues where the NAS Committee left off and fill in the gaps so that *Toxicity Testing in the 21st Century* can actually become influential on regulatory authorities worldwide.

Washington, DC, USA.

*Address correspondence to E. Donald Elliott, Yale Law School, New Haven, CT, USA; e.donald.elliott@yale.edu.

[1] *Daubert v. Merrell Dow Pharmaceuticals, Inc.*, 509 U.S. 579 (1993).

[2] Bret C. Cohen, *Toxicity testing in the 21st century: Better results, less use of animals, legal obstacles are bumps, not roadblocks. Environmental Forum*, March–April 2008.

Risk Analysis, Vol. 29, No. 4, 2009 DOI: 10.1111/j.1539-6924.2008.01167.x

Commentary

High-Throughput Testing—The NRC Vision, The Challenge of Modeling Dynamic Changes in Biological Systems, and the Reality of Low-Throughput Environmental Health Decision Making

Dale Hattis*

An NRC report *Toxicity Testing in the 21st Century* calls for a long-term project to replace much *in vivo* animal testing for "apical" endpoints (end results of pathological processes). It would substitute a large ensemble of high-throughput *in vitro* test systems involving short-term changes in gene expression intended to identify concentrations of environmental chemicals that sufficiently perturb particular toxicity pathways (in tissue culture cells derived from humans) to be of concern. This proposal would direct major future toxicity testing resources to inform choices on the use of many poorly studied chemicals within the traditional no-effect-level paradigm, and without quantitative assessment of benefits as reductions in health risks and associated uncertainties.

By contrast, a competing revolutionary proposal would expand the reach of quantitative assessment of health risks. This would involve replacing the current set of safety/uncertainty factors for noncancer risk assessment with distributions based on empirical data, and quantitative assessment of likely interactions of chemical exposures with background processes involved in human pathological conditions. Rather than a vision of static homeostatic systems, it emphasizes analysis of dynamic changes in protective feedback systems, including errors in initial set up and eventual degradation of homeostasis with ageing. Quantitative treatment of uncertainties in this alternative paradigm would facilitate "value of information" analysis for addition of specific types of tests for clarifying consequences of alternative regulatory options for health protection.

Clark University, St. Louis, MO, USA

*Address correspondence to Dale Hattis, Clark University, 950 Main St., Worcester, MA 01610, USA; tel: 617-283-2521; fax: 508-751-4600; dhattis@clarku.edu.

1. THE INTELLECTUAL CHALLENGE OF ELUCIDATING REAL CHANGES IN BIOLOGICAL FEEDBACK CONTROL SYSTEMS

The NRC recommendations are based on a limited traditional view of biological systems as in unchanging stable equilibria—requiring a "perturbation" of some significant magnitude in order to induce significant effects (and, therefore, the general existence of threshold doses/concentrations below which no functionally important changes are expected). In fact, however, biological systems are generally in a state of cyclic flux in several different modes on different time scales. Additionally, tissue cultures based on single cell types from single people are unlikely to adequately replicate the diversity of sensitivities and modes of response of the pulsating intercellular feedback control systems present *in vivo* in the diverse human population.

As those of us of the baby boomer generation can personally attest, homeostatic controls degrade with aging, leading to both progressive losses of function, and increases in vulnerability to a host of perturbations from external influences. Our "homeostatic" equilibria are at best "meta-stable" states where, like a vortex gurgling down a drain, the appearance of stability can only be temporary and dependent on a continual input of energy. Many of the conditions that cause human mortality and morbidity are likely to be the results of chronic cumulative losses of components of homeostatic control systems.

Each homeostatic feedback system has three kinds of elements: (1) sensors to monitor key parameters, (2) signaling to transmit needs for adaptive responses, and (3) effector mechanisms to reduce detected departures from set points. Systems

of this sort must be set up to operate on different scales of time, physical distance, and levels of biological organization. Key challenges for biological science in the coming decades include how the set points are actually set and how the degree of vigor in response to violations of the set points is determined. Processes that change the set points in response to different circumstances[1] also need to be understood as well as how those changes affect system susceptibility to perturbations. To this observer it seems likely that rather than having the set points directly coded in the genome, a more likely and interesting idea is that the genome somehow encodes a program for the system to "learn" where to best set the set points and how to best respond to various degrees of departure.

Most generally, homeostatic capacities, set points, and control response patterns should be seen as the products of a series of compromises the organism makes to achieve multiple objectives with limited resources. Limitations on those resources, particularly during development, leads to increased risks both in infancy and with aging—increased infant mortality and increased type 2 diabetes in adults are both associated with fetal growth restriction.[2,3] This leads to an entirely different set of expectations for at least some dose-response relationships than would be presumed under the traditional toxicological paradigm.

2. POTENTIAL EFFECTS OF HIGH-THROUGHPUT TESTING IN THE CONTEXT OF LOW-THROUGHPUT ENVIRONMENTAL HEALTH DECISION MAKING

High-throughput testing may ultimately be of substantial value for the large number of new agents that are introduced each year and for many agents introduced in the past with little or no traditional toxicity testing. However, for higher-profile decisions on major agents in commerce involving substantial tradeoffs of risks and economic impacts for different policy options one may doubt that the results of high-throughput tests will ever be sufficient. In this "risk context," at least based on the current EPA regulatory framework, high-throughput test results are much more likely to muddy the waters by raising possibilities for alternative modes of action. Such issues may only be resolved by extensive *in vivo* testing using knock-out mice for specific genes and other high-tech, expensive, and slow test systems. Thus both the potential for savings in testing resources and ultimate contributions for decision making from high-throughput testing seem doubtful in the context of increasing demands for quantification of likely health benefits prior to undertaking costly high-profile regulatory choices. By contrast, a system that attempts to quantitatively assess both actual health risks and the large associated uncertainties[4] can provide relevant input for policymakers and enable analysis of the expected "value of information" benefits of successively adding different types and amounts of information. This has recently been illustrated for pharmacokinetic/pharmacodynamic interindividual variability data.[5]

REFERENCES

1. Schulkin J. Rethinking homeostasis—Allostatic Regulation in Physiology and Pathophysiology. Cambridge, MA: MIT Press, 2003.
2. Forsen T, Eriksson J, Tuomilehto J, Runanen A, Osmond C, Barker D. The fetal and childhood growth of persons who develop type 2 diabetes. Annales Internae Medicinae, 2000; 133:176-182.
3. Hattis D. Strategies for assessing human variability in susceptibility, and using variability to infer human risks. Pp. 27-57 in Neumann DA, Kimmel CA (eds). Human Variability in Response to Chemical Exposure: Measures, Modeling, and Risk Assessment. Boca Raton, FL: CRC Press, 1998.
4. Hattis D, Baird S, Goble R. A straw man proposal for a quantitative definition of the RfD, in Final Technical Report, U.S. Environmental Protection Agency STAR Grant # R825360, "Human Variability in Parameters Potentially Related to Susceptibility for Noncancer Risks." Available at: http://www2.clarku.edu/faculty/dhattis; shortened version in Drug and Chemical Toxicology, 2002; 25:403-436.
5. Hattis D, Lynch MK. Empirically observed distributions of pharmacokinetic and pharmacodynamic variability in humans—Implications for the derivation of single point component uncertainty factors providing equivalent protection as existing RfDs. Pp. 69-93 in Lipscomb JC, Ohanian EV (eds). Toxicokinetics in Risk Assessment. Informa Healthcare, USA, Inc., 2007.

Risk Analysis, Vol. 29, No. 4, 2009 DOI: 10.1111/j.1539-6924.2008.01168.x

Commentary

Toxicity Testing in the 21st Century: Implications for Human Health Risk Assessment

Robert J. Kavlock,[1]* Christopher P. Austin,[2] and Raymond R. Tice[3]

The risk analysis perspective by Daniel Krewski and colleagues lays out the long-term vision and strategic plan developed by a National Research Council committee,(1) sponsored by the U.S. Environmental Protection Agency (EPA) with support from the U.S. National Toxicology Program (NTP), to "advance the practices of toxicity testing and human health assessment of environmental agents." Components of the vision include chemical characterization; the use of human-cell-based, high-throughput assays that cover the diversity of toxicity pathways; targeted testing using animals to fill in data gaps; dose-response and extrapolation modeling; and the generation and use of population-based and human exposure data for interpreting the results of toxicity tests. The strategic plan recognizes that meeting this vision will require a major research effort conducted over a period of a decade or more to identify all of the important toxicity pathways, and that a clear distinction must be made between which pathway perturbations are truly adverse (i.e., would likely lead to adverse health outcomes in humans) and those that are not. Krewski *et al.* note that achieving this vision in a reasonable timeframe (i.e., decades) would require the involvement of an interdisciplinary research institute that would be coordinated and funded primarily by the U.S. federal government and that would foster appropriate intramural and extramural research. It is expected that this approach would greatly increase the number of compounds that can be tested, while providing data more directly relevant for conducting human health risk assessment. The NTP though its Roadmap,[4] the National Institutes of Health (NIH) Chemical Genomics Center (NCGC) through the Molecular Libraries Initiative,[5] and the EPA through its ToxCast™ program[6] and its draft Strategic Plan for the Future of Toxicity Testing have individually recognized the need to bring innovation into the assessment of the toxicological activity of chemicals, and each has made progress in doing so. However, the grand challenge put forth by Krewski *et al.* requires an effort unparalleled in the field of toxicology and risk assessment.

In recognition of the importance of the NRC report(1) and to accelerate progress in this area, two NIH institutes and EPA have entered into a formal collaboration known as Tox21 to identify mechanisms of chemically induced biological activity, prioritize chemicals for more extensive toxicological evaluation, and develop more predictive models of *in vivo* biological response.(2) Consistent with the vision outlined by Krewski *et al.*, success in achieving these goals is expected to result in methods for toxicity testing that are more scientific and cost effective as well as models for risk assessment that are more mechanistically based. As a consequence, a reduction or replacement of animals in regulatory testing is anticipated to occur in parallel with an increased ability to evaluate the large numbers of chemicals that

[1]National Center for Computational Toxicology, Office of Research and Development, U.S. Environmental Protection Agency, Research Triangle Park, NC, USA.
[2]NIH Chemical Genomics Center, National Human Genome Research Institute, National Institutes of Health, MSC 3370, Bethesda, MD, USA.
[3]Biomolecular Screening Branch, National Toxicology Program, National Institute of Environmental Health Sciences, Research Triangle Park, NC, USA.
*Address correspondence to Robert J. Kavlock, Director, National Center for Computational Toxicology, Office of Research and Development, U.S. Environmental Protection Agency, Research Triangle Park, NC 27711, USA; tel: 919-541-2326; fax: 919-541-1194; Kavlock.Robert@epamail.epa.gov.

[4] Available at: http://ntp.niehs.nih.gov/go/vision.
[5] Available at: http://www.ncgc.nih.gov/.
[6] Available at: epa.gov/ncct/toxcat.

currently lack adequate toxicological evaluation. Ultimately, Tox21 is expected to deliver biological activity profiles that are predictive of *in vivo* toxicities for the thousands of understudied substances of concern to regulatory authorities in the United States, as well as in many other countries.

The Tox21 collaboration is being coordinated through a five-year Memorandum of Understanding (MoU),[7] which leverages the strengths of each organization. The MoU builds on the experimental toxicology expertise at the NTP, headquartered at the NIH National Institute of Environmental Health Sciences (NIEHS); the high-throughput screening (HTS) technology of the NIH Chemical Genomics Center (NCGC), managed by the National Human Genome Research Institute (NHGRI); and the computational toxicology capabilities of the EPA's National Center for Computational Toxicology (NCCT). Each party brings complementary expertise to bear on the application of novel methodologies to evaluate large numbers of chemicals for their potential to interact with the myriad of biological processes relevant to toxicity. A central aspect of Tox21 is the unique capabilities of the NCGCs high-speed, automated screening robots to simultaneously test thousands of potentially toxic compounds in biochemical and cell-based HTS assays, and an ability to target this resource toward environmental health issues. As mentioned by Krewski *et al.*, EPA's ToxCast™ Program is an integral and critical component for achieving the Tox21 goals laid out in the MoU.

To support the goals of Tox21, four focus groups—Chemical Selection, Biological Pathways/Assays, Informatics, and Targeted Testing—have been established; these focus groups represent the different components of the NRC vision described by Krewski *et al.* The Chemical Selection group is coordinating the selection of chemicals for the Tox21 compound library to test at the NCGC. A chemical library of nearly 2,400 chemicals selected by NTP and the EPA is already under study at the NCGC and results from several dozen HTS assays are already available. In the near term, this library will be expanded to approximately 8,400 compounds, with an additional ~1,400 compounds selected by the NTP, ~1,400 compounds selected by the EPA, and ~2,800 clinically approved drugs selected by the NCGC. Compound selection is currently based largely on the compound having a defined chemical structure and known purity; on the extent of its solubility and stability in dimethyl sulfoxide (DMSO), the preferred solvent for HTS assays conducted at the NCGC; and on the compound having low volatility. Implementing quality control procedures for ensuring identify, purity, and stability of all compounds in the library is an important responsibility of this group. A subset of the Tox21 chemical library will be included in Phase II of the ToxCast program, which will examine a broader suite of assays in order to evaluate the predictive power of bioactivity signatures derived in Phase I. Phase II of ToxCast will be launched by the summer of 2009.

The Biological Pathways and Assays group is identifying critical cellular toxicity pathways for interrogation using biochemical- and cell-based high-throughput screens and prioritizing HTS assays for use at the NCGC. Assays already performed at the NCGC include those to assess (1) cytotoxicity and activation of caspases in a number of human and rodent cell types, (2) up-regulation of p53, (3) agonist/antagonist activity for a number of nuclear receptors, and (4) differential cytotoxicity in several cell lines associated with an inability to repair various classes of DNA damage. Other assays under consideration include those for a variety of physiologically important molecular pathways (e.g., cellular stress responses) as well as methods for integrating human and rodent hepatic metabolic activation into reporter gene assays. Based on the results obtained, this group will construct test batteries useful for identifying hazard for humans and for prioritizing chemicals for further, more in-depth evaluation.

The Informatics group is developing databases to store all Tox21-related data and evaluating the results obtained from testing conducted at the NCGC and via ToxCast™ for predictive toxicity patterns. To encourage independent evaluations and/or analyses of the Tox21 test results, all data as well as the comparative animal and human data, where available, will be made publicly accessible via various databases, including EPAs Aggregated Computational Toxicology Resource (ACToR), NIEHS' Chemical Effects in Biological Systems (CEBS), and the National Center for Biotechnology Information's PubChem.

As HTS data on compounds with inadequate testing for toxicity becomes available via Tox21, there will be a need to test selected compounds in more comprehensive assays. The Targeted Testing group is developing strategies and capabilities for this purpose using assays that involve higher order testing

[7] Available at: http://ntp.niehs.nih.gov/go/28213.

systems (e.g., roundworms (*Caenorhabditis elegans*), zebrafish embryos, rodents).

In addition to the testing activities, the MoU promotes coordination and sponsorship of workshops, symposia, and seminars to educate the various stakeholder groups, including regulatory scientists and the public, with regard to Tox21-related activities. Persons interested in following the progress of Tox21 are invited to join the EPA's Chemical Prioritization Community of Practice,[8] which meets monthly via teleconference.

Given the scope of the challenge presented by Krewski *et al.*, success will require a long-term concerted effort by a large number of investigators, working in a coordinated manner. The Tox21 consortium welcomes participation in our effort by individual scientists and by organizations. The implications for success of this effort are considerable. If successful, we will be able to address regulatory demands such as those placed by the Food Quality Protection Act for the endocrine screening program[9] and the new European Community Regulation on chemicals and their safe use, known as REACH (Registration, Evaluation, Authorization and Restriction of Chemical Substances),[10] identify key modes of action on a scale not imaginable even a few years ago, direct a much more efficient and effective use of animals in toxicity testing, identify potentially susceptible subpopulations based on the presence of polymorphisms in toxicity pathways, screen the effects of mixtures, and study emerging issues like the safety of nanomaterials. The acquisition of data from broad-scale HTS programs also creates demands to integrate this knowledge and understand the implications for systems biology, and to have risk assessors trained and conversant in the new technologies and their utilities. While the ultimate goal of eliminating the use of animals in toxicology testing might seem unattainable, it is only by carefully evaluating the relevance and reliability of strategies based on *in vitro* test methods that the utility and limitations of such an approach can be determined and decisions made on how best to conduct toxicology testing in the future. To do otherwise will result in increasing demands being placed on systems never designed to handle the large numbers of chemicals in need of evaluation, and continued reliance on test methods based on empirical observation rather than on mechanistic understanding.

[8] Available at: http://www.epa.gov/ncct/practice_community/category_priority.html.

[9] Available at: http://www.epa.gov/opp00001/regulating/laws/fqpa/.

[10] Available at: http://ec.europa.eu/environment/chemicals/reach/reach_intro.htm.

DISCLAIMER

The research described in this report has been funded by one or more of the participating federal agencies. The report does not necessarily reflect the views of the respective organizations.

REFERENCES

1. National Research Council. Toxicity Testing in the 21st Century: A Vision and A Strategy. Washington, DC: National Academy Press, 2007.
2. Collins FS, Gray GM, Bucher JR. Transforming environmental health protection. Science, 2008: 319:906-907.

Risk Analysis, Vol. 29, No. 4, 2009 DOI: 10.1111/j.1539-6924.2008.01169.x

Commentary

Risk Assessment in the 21st Century: Changes Wrought by Changing Science

Lorenz R. Rhomberg*

Krewski *et al*. (2009) have provided an admirably concise overview of the NRC (2007) report, *Toxicity Testing in the 21st Century: A Vision and a Strategy*. The full report is rich with insights and ideas, and I would urge readers of *Risk Analysis* to study its fuller development of the questions addressed by Krewski *et al*. The NRC Committee envisions not simply the advent of new testing technologies (although these will come) but a consequent shift in the questions that toxicology is able to address and therefore will seek to elucidate. The focus will shift away from apical endpoints in whole animals (with subsequent work attempting to explain underlying causes and to assess their potential operation in humans at environmental exposure levels) and toward detecting perturbation of normal physiology itself as the primary object of testing (with subsequent inferences about the disease processes in humans that might be affected). That is, the flow of inference is being reversed—no longer from ultimate effects backward to potential underlying causes but now from causes *forward* to their ultimate potential effects. The NRC report argues that this would be more efficient and better able to address concerns such as human relevance, sensitive subpopulations, and interindividual variability. It would be a mistake to regard testing for physiological perturbation as only the development of better biomarkers or predictors of traditional apical endpoints—the perturbations themselves are the object of the envisioned new paradigm in testing.

Whether, how quickly, and by what means the NRC's vision can be realized is a worthy topic of discussion. The vision supposes that our qualitative and quantitative understanding of the projected consequences of measured perturbations will be sufficiently robust to allow risk assessors, with sufficient confidence, to identify exposure levels that should be expected to be without adverse consequences among members of the exposed population. This requires a considerable ability to identify all potentially relevant toxicity pathways and to model their sequellae and interactions. For this inference to be foolproof will require considerable development of our understanding of biological systems and their behaviors under stress, but the proper comparison is to our current far-from-perfect system in which high-dose animal responses are produced, forcing risk assessment to contend with often unanswerable questions regarding the relevance of these outcomes to human physiology and the potential actions of the toxicant at low exposures. Uncertainties will remain, but they will be different from those uncertainties we have grappled with up to now, and at first they will be less familiar.

How will risk assessment have to change to deal with the new testing approaches imagined by the NRC Committee? Risk assessment and testing are inevitably intertwined, and we need to contend with the mismatch between what practical experiments can reliably find out and what risk assessors (and risk managers) wish to know about the risk or safety of low doses of chemicals to humans. Risk assessment seeks a scientific basis for public health decisions (guidance, if not firm answers). It must deal with the incompleteness, extrapolation uncertainty, and contradictions inherent in experimental animal data when applied for this purpose. The NRC proposal envisions some questions of interest being addressed by testing more directly than they are now (the identity and action of physiological defenses and the time course of their contention with chemically induced stresses, biological variation and its impact

Gradient Corporation, Cambridge, MA, USA.

*Address correspondence to Lorenz R. Rhomberg, Gradient Corporation, 20 University Road, Cambridge, MA 02138, USA; tel: +1-617-395-5552; fax: +1-617-395-5001; lrhomberg@gradientcorp.com.

on sensitivity, the role and impact of physiological differences in test species and in humans) and other questions being addressed less directly (the generation of frank *in vivo* adverse outcomes, the avoidance of which is the ultimate aim of toxicology and risk analysis).

Clearly, our rules of inference and approaches to weight of evidence will have to evolve to reflect toxicity testing that is aimed at providing insight into different questions and makes for different pathways to the answers we seek. Structures such as the human-relevance/mode-of-action framework will be critical, but they will serve not as retrospective assessments of the plausibility of inferred human relevance of apical *in vivo* endpoints in test animals but rather as frameworks for the interpretation of the consequences of observed physiological perturbations. Physiologically based pharmacokinetic modeling will be needed to relate tissue-level effects (which will now be primary data rather than ancillary observations) to the doses and dose patterns that engender them. Systems biology approaches to pharmacodynamics will be needed to relate stresses to perturbations and to determine the degree of perturbation necessary to exceed accommodation and produce frankly adverse effects. Variations in biology will need to be interpreted as to their impact on sensitivity of the whole living system to adverse alteration, and the roles of multiple stressors and their interacting effects on perturbations will need to be addressed.

Risk assessment will still seek to identify no-effect levels, but these will have to be redefined not as the absence of frank impacts on apical endpoints, but as insufficient perturbation of normal processes to cause concern for such impacts. It would be a setback if we simply treated perturbations as more sensitive "precursors" and sought doses without measurable impacts of any sort (on the grounds that they "might have" adverse consequences); the central question of risk analysis must become the identification of doses that cause a failure of defensive and accommodative changes, thereby setting off a cascade of more consequential failures that eventually produce frank toxicity at the whole-organism level. We must be able to make these inferences in a negative way (i.e., being confident that doses not causing excessive perturbations among a finite set of tested toxicity pathways will be unlikely to cause toxicity of any kind of interest for human health protection) and in a positive way (i.e., correctly interpreting how larger perturbations, perhaps of several sorts from several chemicals and varying in intensity among individuals with different genetic backgrounds, histories, and lifestyles, interact and develop over time, perhaps in the face of fluctuating dose levels, to produce unacceptable and adverse reactions).

In truth, we are part of the way along this evolutionary path already. Up until about the mid-1980s, the main question in weight-of-evidence evaluation was to assess whether reliable, repeatable animal responses had been observed—that is, to unmask potential false positive bioassays—with the presumption that any true animal effects were indicative of a similar process in humans. With the progressive elucidation of underlying modes of action, and the realization that many effects operate rather differently in different species or at different dose levels, it has become progressively more important to approach inferences about human risk potential by assessing (rather than presuming) the commonality of underlying causative pathways in humans and animals, and at high doses and low ones. The focus in recent years on sensitive subpopulations and inter-human variability has only increased the need to examine how quantitative variation in underlying influences play out to affect the probability and magnitude of adverse reactions. Already, the bulk of risk evaluation concerns itself with how to address these underlying factors, often without much scientific basis for the discussion. The change in focus of toxicity testing as proposed by the NRC Committee—away from effects and toward causes—should facilitate progress along these lines. But we should be aware that risk assessment, as well as toxicity testing, must evolve.

Risk Analysis, Vol. 29, No. 4, 2009 DOI: 10.1111/j.1539-6924.2008.01170.x

Commentary

Advances in Toxicity Testing Herald Improvements and Challenges for Risk Assessment

Joyce S. Tsuji* and Michael R. Garry

Advancements in quantifying the action of substances on the molecular level combined with increased knowledge on the influence of genetic variation are heralding a new era of toxicological methods to support more accurate assessments of health risk for the human population. The new tools show promise of leading us away from administering high doses to animals and counting bodies, pathological lesions, or tumors (or assessing disease incidence in exposed human populations), toward investigating the mechanistic action and effects of chemicals down to very low levels.

As summarized by Krewski *et al.*,[1] the potential advantages of such approaches are clear, including less animal testing, rapid results using *in vitro* methods, greater throughput for evaluating the enormous number of chemicals in production, and increased accuracy and representativeness for the human population. Such methods would greatly benefit safety assessments of emerging technologies such as new "miracle" products, drugs, and devices containing nanomaterials that desperately need effective rapid screening tests that do not require lifetime animal studies or epidemiological studies. In addition, such methods may allow measurement of potential health risks at relevant exposure levels, which likewise are difficult with current animal testing or epidemiological methods alone because of limited statistical power and ability to detect subtle changes at low doses.

Krewski *et al.*[1] also note that such new approaches will not be without challenges and will likely require considerable validation efforts. And while the entire risk assessment paradigm may not need overhaul to incorporate these approaches, certainly new approaches for conducting risk assessment within that paradigm and in risk management decision making will be required. However, the use of *in-vitro*-derived mechanistic data in place of extensive *in vivo* studies to characterize hazard for classes of chemicals is not foreign to current approaches. One need look no further than risk assessment approaches for dioxins and dioxin-like compounds, where standard *in vitro* experimental designs have been outlined and applied for developing toxic equivalency factors and, thus, predicting relative potency of congeners for which extensive *in vivo* data are not available.[2]

The difficulty in developing *in vitro* assays that predict *in vivo* outcomes with adequate sensitivity and specificity will continue to be a challenge. Recent evaluations of the ability of *in vitro* tests to predict pulmonary toxicity *in vivo* for nanoscale particles (carbonyl iron, crystalline and precipitated amorphous silica, and zinc oxide) have so far not met with much success or confidence.[3] In this study, a key source of discrepancy between *in vitro* and *in vivo* toxicity for zinc oxide was likely the ability of physiological systems to regulate levels of essential elements such as zinc, despite its high reactivity in cell cultures.

Thus, although the mechanism of action of substances occurs on a molecular level, the result of such perturbations is on a whole-organism level. In another example, genetic polymorphisms in the CYP2E1 gene result in a four-fold difference in enzyme expression and a six-fold difference in oxidized trichloroethylene metabolites, but may ultimately result in only 2% or less variability in oxidized trichloroethylene metabolites between the 5th

Exponent, Bellevue, WA, USA.

*Address correspondence to Joyce S. Tsuji, Principal, Exponent, 15375 SE 30th Place, Suite 250, Bellevue, WA 98007, USA; tel: 425-519-8768; fax: 425-519-8799; tsujij@exponent.com.

and 95th percentile in the population because liver perfusion rate, which is not as variable, limits the full expression of this variation.[4,5] Pharmacokinetic analyses, biomonitoring, and other evidence are thus necessary parts of the vision to evaluate the effect of cellular or pathway perturbations on a whole-organism level.[1]

A key challenge for regulatory and public health decision making will be how to define the point of departure for a significant effect. The established method for developing toxicity criteria for noncancer effects of chemicals has been to select a lowest observable adverse effect level (LOAEL) or no observable adverse effect level (NOAEL) from studies in animals or humans and then add factors to lower the dose, depending on various database uncertainties, including extrapolation from a LOAEL to a NOAEL, effects in animals to humans, intra-individual sensitivity, or database deficiencies. In this default paradigm, ignorance is bliss, with the assumption that no effects occurred in the test animals below the NOAEL dose and that once factors were added for the other potential sources of uncertainty, the resulting reference dose would be protective of health. More recent uses of statistical techniques such as benchmark dose analysis help capture some of the statistical and model uncertainty related to the number of animals and variation in response, although these improved approaches are still limited to some extent by dose selection and the endpoints measured.

A result of more refined testing on a cellular level is that more chemicals could have no discernable thresholds for effects. Therefore, considerable evaluation will need to focus on identifying significant perturbations on a molecular level that would result or contribute to adverse consequences to an organism. Practical thresholds likely exist particularly for chemicals that are produced endogenously in the body (e.g., formaldehyde, acetaldyde, hydrogen sulfide) or are present at background levels because of natural occurrence (e.g., arsenic), or are nutritionally essential (e.g., selenium, magnesium, copper, zinc). Such information will need to be considered in more evidence-based approaches for determining points of departure in setting toxicity criteria.

From past experience, it is clear that developing risk assessment toxicity criteria (e.g., reference doses or cancer slope factors), even for more data-rich chemicals (e.g., arsenic) with considerable epidemiological and mechanistic data, is not necessarily clearer or easier than for chemicals with less information. The refined testing methods and approaches will require that public health decision making consider more evidence-based approaches for dealing with uncertainty. As our understanding of toxicity mechanisms increases, so does our awareness of what we don't know.

> As the circle of light increases, so does the circumference of darkness around it.
>
> Einstein

However, as the circle of light increases, the ratio of light to darkness also increases. So while we may be increasingly aware of all we don't know, at least we are moving toward toxicity assessments that are guided by more knowledge and less uncertainty.

REFERENCES

1. Krewski D, Andersen ME, Mantus E, Zeise L. Toxicity testing in the 21st century: Implications for human health risk assessment. Risk Analysis 2009;
2. Van Den Berg M, Birnbaum LS, Denison M, De Vito M, Farland W, Feeley M, Fiedler H, Hakansson H, Hanberg A, Haws L, Rose M, Safe S, Schrenk D, Tohyama C, Tritscher A, Tuomisto J, Tysklind M, Walker N, Peterson RE. The 2005 World Health Organization reevaluation of human and mammalian toxic equivalency factors for dioxins and dioxin-like compounds. Toxicological Science 2006; 93:223-241.
3. Sayes CM, Reed KL, Warheit DB. Assessing toxicity of fine and nanoparticles: Comparing *in vitro* measurements to *in vivo* pulmonary toxicity profiles. Toxicological Science 2007; 97:163-180.
4. Lipscomb JC, Keddaris GL. Use of Physiologically Based Pharmacokinetic Models to Quantify the Impact of Human Age and Interindividual Differences. Washington, DC: U.S. Environmental Protection Agency. EPA/600/R-06/014A, 2006.
5. Lipscomb JC, Teuschler LK, Swartout J, Popken D, Cox T, Kedderis GL. The impact of cytochrome P450 2E1-dependent metabolic variance on a risk-relevant pharmacokinetic outcome in humans. Risk Analysis 2003; 23:1221-1238.

Risk Analysis, Vol. 29, No. 4, 2009 DOI: 10.1111/j.1539-6924.2009.01218.x

Response to Commentaries

Reply to Invited Commentaries on Toxicity Testing in the 21st Century: Implications for Human Health Risk Assessment

Daniel Krewski,[1]* Melvin E. Andersen,[2] Ellen Mantus,[3] and Lauren Zeise[4]

We are grateful for the detailed and thoughtful comments in the six commentaries that accompany our perspective[1] on the NRC report[2] *Toxicity Testing in the 21st Century: A Vision and a Strategy*. The report argues for a pervasive, transformative overhaul of both toxicity testing and the interpretive tools for assessing the relevance of toxicity-test results for humans exposed to potentially toxic chemicals at environmental levels. As might be expected for a report with such broad implications for chemical risk assessment, the commentaries touch on a wide range of salient issues. Rather than discussing every point in detail in our response, we encourage readers of *Risk Analysis* to use our perspective and the accompanying commentaries as a starting point for a more complete evaluation of the future directions for toxicity as set out in the full NRC report.[2] We do, however, welcome the opportunity to highlight one or two points from each of the commentaries here.

Tsuji and Garry[3] note challenges in population health risk assessment[4] based on *in vitro* test results and comment on the need for evidence-based approaches for dealing with uncertainty. Understanding toxicity pathways and the development of high-throughput *in vitro* assays to identify toxicologically important pathway perturbations are emphasized in the NRC vision for the future of toxicity testing. This will require a substantive investment of resources and intellectual capital and take place over a period of time. As a consequence, the NRC report[2] could not propose specific methodologies for using measures of toxicity-pathway perturbations to develop human exposure guidelines. Risk assessment methods based on pathway perturbations will be guided by progress in understanding the dose-response characteristics of such perturbations, rather than simply by reliance on a mandated suite of uncertainty factors. Dose-response models for pathway perturbations will also have to take into account the relationship between pathway perturbations and downstream integrated responses at the cellular or organ level, and how these may vary among individuals: our understanding of such relationships will be greatly enhanced through the mapping of key toxicity pathways.

Tsuji and Garry's[3] reference to chemicals produced endogenously, or present at significant background levels, highlights the need to consider the interplay of multiple exposures and interindividual variability in response. As discussed in Chapter 5 of the NRC report,[2] a key research issue deals with the effect of adding small amounts of toxicants to a toxicity pathway, in light of the mix of endogenous and exogenous exposures that humans experience. Dose-dependent transitions in moving from lower dose levels at which cellular response pathways are not activated to higher dose levels at which activation leading to overt toxicity can be expected to occur need to be understood in the context of underlying human exposures, disease processes, and

[1] McLaughlin Centre for Population Health Risk Assessment, University of Ottawa, Ottawa, Ontario, Canada.
[2] Hamner Institutes for Health Research, Research Triangle Park, NC, USA.
[3] National Research Council, Washington, DC, USA.
[4] California Environmental Protection Agency, Oakland, CA, USA.
* Address correspondence to Daniel Krewski, McLaughlin Centre for Population Health Risk Assessment, University of Ottawa, Ottawa, Ontario K1N 6N5, Canada; tel: 613-562-5381; fax: 613-562-5380; dkrewski@uottawa.ca; cphra@uottawa.ca.

health status. This is central to improvements in risk assessment, as discussed in the recent NRC report *Science and Decisions: Advancing Risk Assessment.*[5] The lessons learned from endogenous compounds demonstrating toxicity should help guide the development of new approaches for risk assessment and decision making for environmental agents based on *in vitro* test results. Dose-response models that take into account both the pharmacokinetics and pharmacodynamics of toxicity pathways and an understanding of the underlying biological and health status of the population regarding critical toxicity-pathway components will aid the transition to improved risk assessment. As noted in our committee's report[2](p. 80), a detailed understanding of toxicity-pathway activation in the human population "will be critical in understanding the implications of high-throughput results for the population and for identifying susceptible populations." Population-based studies will likely prove to be an essential component of the new toxicity-testing paradigm and the risk assessment approaches that use the test results. As the molecular components of the cell circuitry underlying toxicity pathways are identified, the biological substrates that determine interindividual susceptibility will become more evident, permitting quantitative description of interindividual variability in risk within specific populations and across the general population.

Conolly[6] suggests that the NRC report[2] fails to discuss the accuracy of risk predictions; he also advocates for the use of more detailed pharmacokinetic models in risk assessment. These pharmacokinetic models endorsed by Conolly would have greater resolution within cells and tissues that could assist in the development of "virtual tissue" models to integrate responses across molecules, pathways, cells, and tissues, eventually allowing simulation of expected toxicity at the organ or organism level. The NRC report[2] indirectly discusses concerns with the accuracy of current approaches at the beginning of Chapter 2 (p. 35):

> Current approaches to toxicity testing rely primarily on observing adverse biologic responses in homogeneous groups of animals exposed to high doses of a test agent. However, the relevance of such animal studies for the assessment of risks to heterogeneous human populations exposed at much lower concentrations has been questioned.

In contrast, toxicity-pathway analysis will evaluate responses directly in cells, cell lines, or tissues, preferably of human origin, across a range of doses. The NRC report[2] does not enumerate the criteria for assessing which perturbations should be regarded as adverse. A wide variety of responses are possible with the suites of toxicity-pathway assays envisioned: those that are relevant for human health risk assessment will need to be distinguished from those that are not. Making this distinction will be guided by the mapping of key toxicity pathways, including the determination of which pathway perturbations, if not prevented, can be expected to lead to adverse health outcomes. Key challenges will be to develop assays that are relevant for responses in representative human cells, and how to interpret findings in cells that are engineered to be hyper-responsive. In addition, interpretive tools for examining pathway circuitry and the manner in which perturbations cascade into integrated cellular responses will be essential in understanding and assessing risk assessment based on toxicity-pathway perturbations, rather than on apical endpoints.

Two aspects of Connolly's[6] call for increased detail in microdosimetry and "virtual tissues" warrant comment. Microdosimetry may be used for *in vitro*-to-*in vivo* extrapolations of toxicity-pathway results, depending on the level of knowledge of the pharmacokinetics, the modes of action at the cellular level, and the detail required in the extrapolation. The virtual tissue is a computational *in silico* model primarily geared to predict toxicity from a variety of input data at the molecular, cellular, organ, and organism level: such research could be a valuable adjunct for safety testing in the pharmaceutical industry.[7] The goal of the NRC report[2] is to create a new toxicity-testing paradigm that identifies perturbations of toxicity pathways and uses physiologically based pharmacokinetic (PBPK) modeling of human exposures, and computational dose-response models for these pathways can be used to identify exposures that are not expected to lead to significant risk in specific populations. Rather than predicting apical responses in animals or humans, the objective is the prediction of significant toxicity-pathway perturbations that have the potential for downstream consequences, manifest as adverse health outcomes.

Elliott[8] highlights legal challenges in bringing this new vision to life within the context of regulatory risk assessment. Chapter 6, *Prerequisites for Implementing the Vision in Regulatory Contexts*, of the NRC report[2] covered some of these same issues. The report acknowledged that the vision raises fundamental questions about the definitions of adversity. The report envisions extensive cross-sector

collaboration to ensure that results from a suite of toxicity tests and analysis from new tools for dose-response and *in vitro*-to-*in vivo* extrapolation can be implemented within a regulatory setting. Clearly, evaluation of the issues by a "more legally sophisticated group or institution," as argued by Elliott,[8] will be valuable; some of the analysis conducted by the Environmental Law Institute[9] on alternatives to animal testing under U.S. toxic and pesticide laws is relevant in this regard.

A fundamental issue to be addressed is the scope of the information needed to reflect human health risks and guide the creation of the risk assessment tools for assessing conditions under which humans might experience significant perturbations of key toxicity pathways. As discussed in Chapter 6 of the NRC report,[2] a key issue will be to establish what constitutes an "adverse" effect, a regulatory trigger for many statutes administered by federal agencies. Implementation of the NRC vision for human risk assessment for environmental chemicals will require coordination of the technical development of toxicity-pathway assays and extrapolation tools and the development of clear and explicit guidelines for their use within the regulatory and legal sectors. It will also require understanding and acceptance by nonscientists. A challenge will be to articulate the fundamental biological concepts and principles underlying the vision sufficiently well so that the advantages of the new approach to toxicity testing and assessment of environmental agents, and the regulatory and scientific validity of the concomitant risk assessment methods, are clearly understood.

Hattis[10] discusses challenges in using high-throughput results for risk decision making "based on the current EPA regulatory framework." In addition, he describes an alternative, stating that: "a system that attempts to quantitatively assess both actual health risks and the large associated uncertainties[11] can provide relevant input for policy makers and enable analysis of the expected 'value of information' benefits of successively adding different types and amounts of information." This important point was also made in the recent NRC *Science and Decisions* report.[5] There are many challenges in moving from current risk assessment practices based on apical endpoints to a new set of practices based on perturbation of toxicity pathways; these issues are directly addressed by Elliott[8] and in Chapter 6 in the NRC report.[2] Approaches based on the value of information could prove quite beneficial as individual tests are developed, test batteries are designed, and test methods accepted for regulatory use replace current methods.

It is important to emphasize that the NRC vision is not an augmentation of current methods. Instead, it recommends their replacement over time as new approaches are developed. Importantly, the suite of tests used to identify toxicity-pathway perturbations is not simply testing to evaluate short-term changes in gene expression, although such tests may well play a role. Rather, these tests represent rapid, *in vitro* evaluations of biological targets of chemicals and the dose-response characteristics of chemical-target interactions. This is the fundamental scientific advantage of these tests for purposes of toxicity testing. The NRC report[2] discusses genomic studies that can be included as part of targeted testing, where shorter-term, multiday dosing studies could include genomic components to assess pathways affected by environmental agents and provide broader evaluation of tissue response than that available from histological analysis alone.

Chapter 5 of the NRC report[2] discusses the implementation of the vision, without making specific proposals regarding suites of pathway tests. The key questions for developing the knowledge to support pathway testing are (box 5-1, p. 124):

- Toxicity-pathway identification—What are the key pathways whose perturbations result in toxicity?
- Multiple pathways—What alterations in response can be expected from simultaneous perturbations of multiple toxicity pathways?
- Adversity—What adverse effects are lined to specific toxicity-pathway perturbations? What patterns and magnitudes of perturbations are predictive of adverse health outcomes?
- Life stages—How can the perturbations of toxicity pathways associated with developmental timing or aging be best captured to enable the advancement of high-throughput assays?
- Effects of exposure duration—How are biologic responses affected by exposures of different duration?
- Low-dose response—What is the effect on a toxicity pathway of adding small amounts of toxicants in light of preexisting endogenous and exogenous human exposures?
- Human variability—How do people differ in their expression of toxicity pathway constituents and in their predisposition to disease and impairment?

With regard to the alternative proposal from Hattis,[10] the NRC committee began its efforts, looking at current test approaches, which are under pressure to meet (and are perhaps not capable of meeting) several competing demands (see Chapter 2, pp. 39–40, and pp. 42–43 of the NRC report[2]):

- the testing of large numbers of existing chemicals, many of which lack basic toxicity data;
- the testing of the large number of new chemicals and novel materials, such as nanomaterials, introduced into commerce each year;
- the evaluation of potential adverse effects with respect to all critical end points and life stages;
- the evaluation of potential toxicity in the most vulnerable members of the human population;
- the use of the fewest animals in the most humane fashion;
- the need to reduce the cost and time required for chemical safety evaluation; and
- the need to develop a more robust scientific basis for risk assessment by providing detailed mechanistic and dosimetric information and by encouraging the integration of toxicological and population-based data.

These demands served as design criteria for developing the vision. The alternative proposal by Hattis[10] needs to be looked at through the lens of these same criteria. Approaches that continue to increase the burden of time-consuming and expensive studies in experimental animals, no matter how sophisticated in relation to uncertainty and variability analysis,[5] do not take full advantage of advances in modern biology that offer the prospect for quantum-level improvements in toxicity testing.

Rhomberg[12] confronts the consequences of the NRC report[2] for risk assessment head-on: he foresees that implementation of the recommendations of the NRC report will require pervasive changes for both toxicity testing and risk assessment. With respect to inference guidelines for risk assessment, Rhomberg[12] comments that: "We must be able to make these inferences in a negative way (i.e., being confident that doses not causing excessive perturbations among a finite set of tested toxicity pathways will be unlikely to cause toxicity of any kind of interest for human health protection)...." Rhomberg's[12] last paragraph provides insights into why a new paradigm is needed and the risk assessment challenges inherent in moving in new directions:

> With the progressive elucidation of underlying modes of action, and the realization that many effects operate rather differently in different species or at different dose levels, it has become progressively more important to approach inferences about human risk potential by assessing (rather than presuming) the commonality of underlying causative pathways in humans and animals, and at high doses and low ones. The focus in recent years on sensitive subpopulations and interhuman variability has only increased the need to examine how quantitative variation in underlying influences play out to affect the probability and magnitude of adverse reactions. Already, the bulk of risk evaluation concerns itself with how to address these underlying factors, often without much scientific basis for the discussion. The change in focus of toxicity testing as proposed by the NRC committee—away from effects and toward causes—should facilitate progress along these lines.

The final commentary by Kavlock *et al.*[13] discusses the convergence of the recommendations of the report with an initiative covered under the memorandum of understanding between three U.S. governmental organizations—the Environmental Protection Agency, the National Chemical Genomics Center, and the National Toxicology Program—announced early in 2008.[14] EPA has since then incorporated many aspects of the NRC vision for the future of toxicity testing into its *Strategic Plan for the Future of Toxicity Testing at the U.S. Environmental Protection Agency*.[15] These initiatives should speed the development of various test approaches and have significant promise for fulfilling a number of the key aspects of the NRC vision.

These federal initiatives focus on "decisions made on how best to construct test batteries for identifying hazards for humans and for prioritizing chemicals for further, more in-depth evaluation." To a certain extent, these initiatives are based on the notion that test methods need to reproduce high-dose results in animals or to support decisions about which tests should be done in animal systems. This latter approach is represented by option 2 in the suite of options evaluated by the NRC committee in its report[2] (cf. table 2-1, p. 44). We appreciate the initiatives of these federal agencies to move incrementally toward the adoption of new methodologies to, in essence, optimize current testing approaches. However, over the long term, the emphasis on prioritization and validation of high-dose animal test results could actually hinder full implementation of the new approach to toxicity testing for the assessment of environmental agents outlined in the NRC report.[2] The emphasis in the NRC report[2] is on human-based test systems, with options 3 and 4 (the intermediate and long-term

options favored by the NRC committee) involving an overhaul of present approaches to evaluate the likelihood that chemicals may influence human health at environmental exposure levels. The NRC committee did not envision that, even in the long term, animal tests could be entirely eliminated. Specifically, targeted testing, which could be *in vivo* or *in vitro*, would be used:

- to clarify substantial uncertainties in the interpretation of toxicity-pathway data;
- to understand effects of representative prototype compounds from classes of materials, such as nanomaterials, that may activate toxicity pathways not included in a standard suite of assays;
- to refine risk estimates when targeted testing can reduce uncertainty, and a more refined estimate is needed for decision making;
- to investigate the production of possibly toxic metabolites of new compounds; and
- to fill data gaps in the toxicity-pathway testing strategy to ensure that critical toxicity pathways and endpoints are adequately covered.

It is time for a coalition of interested parties, including federal agencies, to set a common goal, the fulfillment of the vision elaborated in the NRC report,[2] with specific targets for implementation. As noted previously, the EPA has already taken a major step in this direction through its strategic plan for the future of toxicity testing.[15] The final comment by Kavlock *et al.*[14] emphasizes the need for change:

> While the ultimate goal of eliminating the use of animals in toxicology testing might seem unattainable, it is only by carefully evaluating the relevance and reliability of strategies based on in vitro test methods that the utility and limitations of such an approach can be determined and decisions made on how best to conduct toxicology testing in the future. To do otherwise will result in increasing demands being placed on systems never designed to handle the large numbers of chemicals in need of evaluation, and continued reliance on test methods based on empirical observation rather than on mechanistic understanding.

In conclusion, the authors of this perspective greatly appreciate the opportunity to reflect on the issues raised in the preceding commentaries; these thoughtful comments highlight a range of risk assessment issues that will be confronted in moving to a new toxicity-testing approach that is substantially different from current paradigm. It bears emphasis that the NRC committee believed that the overhaul of current practices is not simply a convenience required to streamline the toxicity-testing process; it is a necessary step for implementing a preferred approach to toxicity testing in the 21st century and achieving the design goals on which the NRC vision rests. These changes are becoming possible with our increasingly sophisticated understanding of human biology, human signaling motifs, and toxicity pathways. How quickly can this transition take place? A progression of important science and technology activities for making a transition to these new methodologies is provided in the NRC report[2] (fig. 5-1, p. 136). Other management approaches could also be envisioned based on available resources and public and political will to bring about these changes; Chapter 5 of the report offers some suggestions in this regard. We very much look forward to a continued dialog within the toxicological and risk assessment communities on key issues in the NRC report and on further discussion of the questions raised by the commentators regarding implementation of the vision.

Finally, the NRC report[2] was completed in 2006, reviewed in late 2006, revised in early 2007, and released on June 12, 2007. Many areas reviewed in relation to the scientific tools and technologies that form the foundation for the NRC vision (see the NRC report,[2] Chapter 4, and Andersen and Krewski,[16] table 1, for further details) have advanced considerably, especially stem cell biology and regenerative medicine and computational systems biology. These advances are important for developing toxicity-pathway assays and dose-response models linking perturbations and more integrated responses. The path forward looks increasingly achievable. The main remaining limitation is finding the political will and resources to move forward with the goals specified in the NRC report.[2] Over the past 18 months, members of the NRC committee have made nearly 50 invited presentations on the vision set out in the NRC report; continued interest in this area is evidenced by future presentations already scheduled throughout 2009.[5] Many of the issues raised by the

[5]In 2007, invited presentations were made to/at: U.S. EPA Assistant Administrator, Annual Summer Meeting of the Toxicology Forum, U.S. EPA Health Effects Research Laboratory, Karolinska Institute for Environmental Medicine (Stockholm, Sweden), Seminar on Aspects on Risk-Benefit Assessment of Food Consumption—Directions for the Future (Uppsala, Sweden), Society for Risk Analysis Annual Meeting, Dow Chemical Company, District of Columbia Bar Association and Environmental Law Institute, and OECD Workshop Integrated

audiences at these presentations have also served to clarify the steps inherent in making such a profound change in toxicity testing and risk assessment. A second perspective[16] regarding the NRC report[2] prepared by two of us, in press in *Toxicological Sciences*, provides complementary materials regarding the toxicity-testing aspects of the report and shows how our thinking has evolved since the NRC report[2] was completed in 2007.

REFERENCES

1. Krewski D, Andersen ME, Mantus E, Zeise L. Toxicity testing in the 21st century: Implications for human health risk assessment. Risk Analysis, 2009; 29(4):474–479.
2. National Research Council. Toxicity Testing in the Twenty-First Century: A Vision and a Strategy. Washington, DC: National Academies Press, 2007.
3. Tsuji JS, Garry MR. Advances in toxicity testing herald improvements and challenges for risk assessment. Risk Analysis, 2009; 29(4):490–491.
4. Krewski D, Hogan V, Turner MC, Zeman PL, McDowell I, Edwards N, Losos J. An integrated framework for risk management and population health. Human and Ecological Risk Assessment, 2007; 13:1288–1312.
5. National Research Council. Science and Decisions: Advancing Risk Assessment. Washington, DC: National Academies Press, 2009.
6. Conolly RB. Commentary on "Toxicity testing in the 21st century: Implications for human health risk assessment" by Krewski et al. Risk Analysis, 2009; 29(4):480–481.
7. Houck KA, Kavlock RJ. Understanding mechanisms of toxicity: Insights from drug discovery research. Toxicology and Applied Pharmacology, 2008; 227:163–178.
8. Elliott ED. Needed: A strategy for implementing the vision. Risk Analysis, 2009; 29(4):482.
9. Environmental Law Institute. Implementing Alternatives to Animal Testing Under U.S. Toxics and Pesticide Laws [online]. Washington, DC: Environmental Law Institute, 2008. Available at: http://www.eli.org/pdf/research/ELI_Implementing_Alternatives_White_Paper_v.1_.pdf, Accessed on January 15, 2009.
10. Hattis D. High-throughput testing—The NRC vision, the challenge of modeling dynamic changes in biological systems, and the reality of low-throughput environmental health decision making. Risk Analysis, 2009; 29(4):483–484.
11. Hattis D, Baird S, Goble R. A straw man proposal for a quantitative definition of the RfD, in Final Technical Report, U.S. Environmental Protection Agency STAR Grant # R825360, "Human variability in parameters potentially related to susceptibility for noncancer risks." Available at: http://www2.clarku.edu/faculty/dhattis, Accessed February 27, 2009; Drug and Chemical Toxicology [shortened version], 2002; 25:403–436.
12. Rhomberg LR. Risk assessment in the 21st century: Changes wrought by changing science. Risk Analysis, 2009; 29(4):488–489.
13. Kavlock RJ, Austin CP, Tice RR. Toxicity testing in 21st century: Implications for human health risk assessment. Risk Analysis, 2009; 29(4):485–487.
14. Collins FS, Gray GM, Bucher JR. Transforming environmental health protection. Science, 2008; 319:906–907.
15. Environmental Protection Agency Future of Toxicity Testing Workgroup. Strategic Plan for the Future of Toxicity Testing at the U.S. Environmental Protection Agency. Washington, DC: Office of the Science Policy Advisor, U.S. Environmental Protection Agency, 2009.
16. Andersen ME, Krewski D. Toxicity testing in the 21st century: Bringing the vision to life. Toxicological Sciences, 2009; 107(2):324–330.

Approaches for Testing and Risk Assessment. In 2008: Risk Assessment Specialty Section of the Society of Toxicology; International Life Sciences Institute Annual Meeting; Health Canada; Workshop on Acute Chemical Safety Testing (joint support from Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) and the National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM)); ICCVAM Ten-Year Anniversary Symposium (Celebrating the Advancement of Public Health and Animal Welfare with Sound Science: Envisioning New Directions in Toxicology); Annual Meeting of the Society for Toxicology; Molecular Medicine Tri-Conference: Bridging Biology, Chemistry, and Business; Centre for Advanced R&D on Alternative Methods, Conference 5: Environment, Health and Safety Testing: Alternative testing, the future for safety testing (Bruges, Belgium); Michigan Regional Society of Toxicology Meeting; Annual Spring Meeting of the Midwest Regional Chapter of the Society of Toxicology; Scientific Advisory Panel of the Institute for In Vitro Sciences; American Chemistry Council Board of Directors; California EPA Mode of Action Workshop: Characterizing Dose Response & Hazard: International Council of Chemical Associations—Long Range Initiative Workshop on Twenty-First Century Approaches to Toxicity Testing, Biomonitoring, and Risk Assessment (Amsterdam, The Netherlands); National Toxicology Program, Scientific Advisory Committee on Alternative Toxicological Methods (SACATM); National Association of Biomedical Research Leadership Conference; California Institute for Regenerative Medicine Workshop on Stem Cells in Predictive Toxicology; 15th International Congress on In Vitro Toxicology (formerly INVITOX) (Stockholm, Sweden); Johns Hopkins University, Center for Alternatives to Animal Testing; Cefic (Conseil Europeen des Federations de l'Industrie Chimique)—The European Chemistry Industry Council Innovative Science for Environment and Health: Tools for Registration, Evaluation and Authorisation of Chemicals (REACH) and Beyond (Brussels, Belgium); Duke University Integrated Toxicology and Environmental Health Program; Unilever Toxicology Science Programme Annual Review (Bedfordshire, United Kingdom); University of Wurzburg (Wurzburg, Germany); Society of Toxicology of Canada; Humane Society of the United States. In 2009: American Association for Advancement of Science, Society for In Vitro Biology Annual Meeting, 30th Anniversary Meeting Netherlands Society of Toxicology (Amsterdam, The Netherlands).