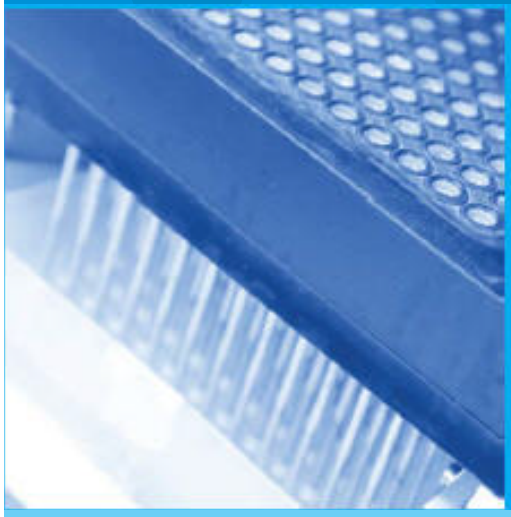
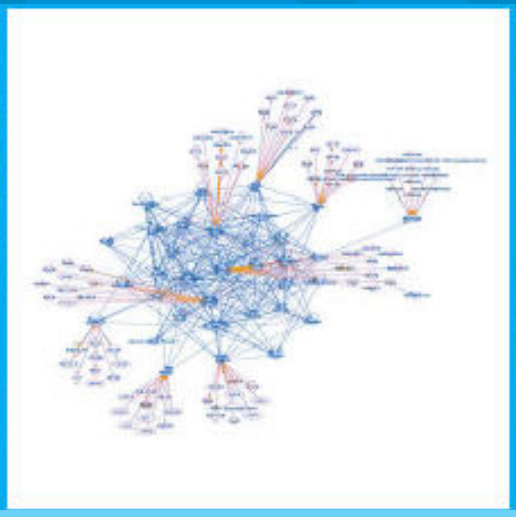
EPA/600/R-14/004 | September 2014 | www.epa.gov/ncea

U ited States
En ironmental Protec ion
Agency

Next Generation Risk Assessment:

Incorporation of Recent Advances in Molecular, Computational, and Systems Biology





Final Report

National Center for Environmental Assessment
Office of Research and Development

United States Environmental Protection Agency

September 2014
EPA/600/R-14/004

Next Generation Risk Assessment:

Recent Advances in Molecular, Computational, and Systems Biology

National Center for Environmental Assessment

Office of Research and Development

U.S. Environmental Protection Agency

Washington, DC 20460

Disclaimer

This document is final. It does not represent and should not be construed to represent any Agency determination or policy. Mention of trade names or commercial products does not constitute endorsement or recommendation for use.

Contents

Figures

Tables

Acronyms and Abbreviations

AC_{50}	concentration at 50 percent of maximum activity
AhR	aryl hydrocarbon receptor
AML	acute myeloid leukemia
AOP	adverse outcome pathway
B[a]P	benzo[a]pyrene
BA	balanced accuracy
BMD	benchmark dose
BMDL	benchmark dose lower limit
BN	Boolean network
CDC	Centers for Disease Control and Prevention
CK	chemokine signaling
CNS	central nervous system
CSS	Chemical Safety for Sustainability
CTD	Comparative Toxicogenomic Database
DNA	deoxyribonucleic acid
ECHA	European Chemicals Agency
ECM	extracellular matrix
EDC	endocrine disrupting chemical
EPA	U.S. Environmental Protection Agency
EWAS	environment-wide association study
GEO	Gene Expression Omnibus
GWAS	genome-wide association study
HC	high-content
HCS	high-content screening
HD	human dose
HHRA	human health risk assessment
HPG	hypothalamus-pituitary-gonad
HPT	hypothalamus-pituitary-thyroid
HT	high-throughput
HTS	high-throughput screening
HTVMD	high-throughput virtual molecular docking
IC_{50}	concentration producing a 50 percent inhibition of response
IC_{10}	concentration producing a 10 percent inhibition of response
IVIVE	*in vitro* to *in vivo* extrapolation
JRC	Joint Research Council
KEGG	Kyoto Encyclopedia of Genes and Genomes
LC_{50}	concentration at 50 percent mortality
LD_{50}	dose at 50 percent mortality
LDA	linear discriminant analysis

Acronym	Abbreviation
LOAEL	lowest observable adverse effect level
MGI	Mouse Genome Informatics
MIE	molecular initiating event
MOA	mode of action
mRNA	messenger ribonucleic acid
NCBI	National Center for Biotechnology Information
NexGen	Next Generation Risk Assessment
NHANES	National Health and Nutrition Examination Survey
NIEHS	National Institute of Environmental Health Sciences
NIH	National Institutes of Health
NRC	National Research Council
NTP	National Toxicology Program
OECD	Organization for Economic Cooperation and Development
ORD	Office of Research and Development
PAH	polycyclic aromatic hydrocarbon
PBPK	physiologically based pharmacokinetic
PBTK	physiologically based toxicokinetic
PCNA	proliferating cell nuclear antigen
PD	pharmacodynamic
PK	pharmacokinetic
POD	point of departure
ppb	part per billion
ppm	part per million
QSAR	quantitative structure-activity relationship
RNA	ribonucleic acid
ROS	reactive oxygen species
RTK	reverse toxicokinetics
SAR	structure-activity relationship
SNP	single nucleotide polymorphism
SOAR	Systematic Omics Analysis Review
TK	toxicokinetic
Tox21	Toxicology in the 21st Century
ToxCast™	Toxicity Forecaster
TSH	thyroid-stimulating hormone
uPAR	plasminogen-activating system
VARIMED	VARiants Informing MEDicine
VEGF	vascular endothelial growth factor
VT	virtual tissue
WHO	World Health Organization

Authors and Acknowledgments

This document reflects contributions of many individuals whom we acknowledge. We also would like to acknowledge Dr. Peter Preuss, who initiated this effort, and Ms. Rebecca Clark, Dr. Kenneth Olden, Dr. John Vandenberg, Ms. Debra Walsh, and Dr. Robert Kavlock for their continued support of this project.

Managing Editors

Ila Cote, Lyle Burgoon, Robert DeWoskin – U.S. Environmental Protection Agency, Office of Research and Development

Authors and Contributors

Executive Summary

Lyle Burgoon,[1] Robert DeWoskin,[1] Elaine Cohen-Hubal,[1] and Ila Cote[1]

Introduction

Ila Cote,[1] Paul Anastas,[2] Stan Barone,[1] Linda Birnbaum,[3] Rebecca Clark,[1] Kathleen Deener,[1] David Dix,[1] Stephen Edwards,[1] and Peter Preuss[1]

Preparation for Prototype Development

Planning for "Fit-for-Purpose" Assessments

Ila Cote,[1] Annie Jarabek,[1] and John Vandenberg[1]

A Framework

Daniel Krewski,[4] Margit Westphal,[4] Greg Paoli,[5] Maxine Croteau,[5] Mustafa Al-Zoughool,[4] Melvin Andersen,[6] Weihsueh Chiu,[1] Lyle Burgoon,[1] and Ila Cote[1]

Science Community and Stakeholder Engagement

Kim Osborn,[7] Gerald Poje,[8] and Ron White[9]

Recurring Issues in Risk Assessment

Daniel Krewski,[4] Melvin Andersen,[6] Kim Boekelheide,[10] Frederic Bois,[11] Lyle Burgoon,[1] Weihsueh Chiu,[1] Michael DeVito,[3] Hisham El-Masri,[1] Lynn Flowers,[1] Michael Goldsmith,[1] Dale Hattis,[21] Derek Knight,[12] Thomas Knudsen,[1] William Lefew,[1] Greg Paoli,[5] Edward Perkins,[13] Ivan Rusyn,[14] Cecilia Tan,[1] Linda Teuschler,[1] Russell Thomas,[1] Maurice Whelan,[15] Timothy Zacharewski,[16] Lauren Zeise,[17] and Ila Cote[1]

Key Questions and Evidence Integration

Ila Cote,[1] Lyle Burgoon,[1] and Robert DeWoskin[1]

The Prototypes

Tier 3: Major-scope Assessments – Benzene-induced Leukemia

Reuben Thomas,[3] Alan Hubbard,[18] Cliona McHale,[18] Luoping Zhang,[18] Stephen Rappaport,[18] Qing Lan,[19] Nathaniel Rothman,[19] Jennifer Jinot,[1] Babasaheb Sonawane,[1] Ila Cote,[1] Martyn Smith,[18] and Kathryn Guyton[28]

Mixtures and Nonchemical Stressors
Timothy Zacharewski,[16] Ila Cote,[1] Linda Teuschler,[1] and Lyle Burgoon[1]

Interspecies Extrapolation
Lyle Burgoon,[1] Ila Cote,[1] and Edward Perkins[13]

Responses at Environmental Exposure Levels
Weihsueh Chiu,[1] Dan Krewski,[4] and Lyle Burgoon[1]

Implications of New Methods for Recurring Issues in Risk Assessment
Ila Cote[1] and Robert DeWoskin[1]

Lessons Learned from Developing the Prototypes
Ila Cote,[1] Robert DeWoskin,[1] Lynn Flowers,[1] John Vandenberg,[1] Douglas Crawford-Brown,[22] and Lyle Burgoon[1]

Challenges and Research Directions
Robert DeWoskin,[1] John Vandenberg,[1] Annie Jarabek,[1] Elaine Cohan-Hubal,[1] Tina Bahadori,[1] and Ila Cote[1]

Other Contributors
Ken Ramos,[23] Peter McClure,[24] Heather Carlson-Lynch,[24] Julie Stickney,[24] Catherine Blake,[25] William Pennie,[26] and Karen Leach[26]

[[1]U.S. Environmental Protection Agency, [2]Yale University, [3]National Institute of Environmental Health Sciences and the National Toxicology Program, [4]University of Ottawa, [5]Risk Sciences International, [6]The Hamner Institutes for Health Sciences, [7]ICF International, [8]Grant Consulting Group, [9]Johns Hopkins University, [10]Brown University, [11]L'Institut National de l'Environnement Industriel et des Risques, [12]European Chemicals Agency, [13]Army Corps of Engineers, [14]University of North Carolina at Chapel Hill, [15]European Joint Research Commission, [16]Michigan State University, [17]California Environmental Protection Agency Office of Environmental Health Hazard Assessment, [18]University of California at Berkeley, [19]National Cancer Institute, [20] formerly at Stanford University, now at Harvard University, [21]Clark University, [22]University of Cambridge, [23]University of Louisville, [24]SRC, [25]University of Illinois, [26]Pfizer, Inc., [27]Oak Ridge Institute for Science and Education, [28]International Agency for Research on Cancer]

Agency Partners

- Army Corps of Engineers: Edward Perkins and Anita Meyer
- California Environmental Protection Agency, Office of Environmental Health Hazard Assessment: George Alexeeff, Martha Sandy, and Lauren Zeise
- Centers for Disease Control, National Center for Environmental Health, and the Agency for Toxic Substances and Disease Registry: Chris Portier (retired), Bruce Fowler (retired), Tom Sinks, and Gwendolyn Price
- Department of Defense, Office of the Secretary of Defense: Robert Boyd and Patrick Mason
- European Chemicals Agency: Derek Knight
- European Joint Research Commission: Maurice Whelan

- FDA's National Center Toxicological Research: Donna Mendrick and William Slikker
- Health Canada: Carole Yauk
- L'Institut National de l'Environnement Industriel et des Risques: Frederic Bois
- National Institutes of Health National Center for Advancing Translational Science: Menghang Xia and Anton Simeonov
- National Institute of Environmental Health Sciences and the National Toxicology Program Linda Birnbaum, Scott Auerbach, John Balbus, Michael DeVito, Elizabeth Maull, Kristine Thayer, and Raymond Tice
- National Institute for Occupational Safety and Health: Christine Sofge, Paul Schulte, Ainsley Weston

U.S. Environmental Protection Agency

- Office of Research and Development: Tina Bahadori, Michael Broder, David Bussard, Vincent Cogliano, Rory Conolly, Dan Costa, John Cowden, Sally Darney, Elizabeth Erwin, Susan Euling. Annette Gatchett, Gary Hatch, Ronald Hines, Elaine Hubal, Annie Jarabek, Maureen Johnson, Robert Kavlock, Channa Keshava, Monica Linnenbrink, Matt Martin, Connie Meacham, Eva McLanahan, Shaun McCullough, David Miller, Kenneth Olden, David Reif, James Samet, Rita Schoeny, Imran Shah, Deborah Segal, Woodrow Setzer, Michael Slimak, John Vandenberg, Rong-Lin Wang, Debra Walsh, John Wambaugh, and Paul White
- Office of Air and Radiation: Souad Benromdhane, Bryan Hubbell, Kelly Rimer, Carl Mazza, Deirdre Murphy, Susan Stone, and Lydia Wegman (retired)
- Office of Chemical Safety and Pollution Prevention: Stan Barone, Vicki Dellarco (retired), Steven Knott, Anna Lowit, Mary Manibusan, Jennifer McLain, Jeff Morris, Anita Pease, Laura Parsons, and Jennifer Seed
- Office of Superfund and Emergency Response: Michele Burgess, Rebecca Clark, Helen Dawson, Stiven Foster, and Kathleen Raffaele
- Office of Water: Cynthia Dougherty, Elizabeth Doyle, and Elizabeth Southerland
- Office of Children's Health Protection: Michael Firestone and Brenda Foos
- Office of Environmental Justice: Charles Lee
- Regional Liaisons: Carole Braverman and Bruce Duncan

Other Individuals and Organizations

- **William Farland (Colorado State University), Julia Gohlke (University of Alabama), Lawrence Lash (Wayne State University), Jennifer McPartland (Environmental Defense Fund), and John Quackenbush (Dana Farber Cancer Institute**) for serving as external peer reviewers and providing helpful feedback
- **Christine Sofge and Paul Schulte (CDC/NIOSH)** for sharing their pre-publication draft manuscript and excellent ideas on advances in risk assessment approaches
- **ICF International** for providing technical support to EPA for this report under contracts EP-C-09-009 and EP-C-14-001

List of Technical Papers in Association with the NexGen Report That Provide Additional Scientific Details

Preparation for Prototype Development	A framework for the next generation of risk science by Daniel Krewski, Margit Westphal, Mel Andersen, Greg Paoli, Weihsueh Chiu, Mustafa Al-Zoughool, Maxine Croteau, Lyle Burgoon, and Ila Cote (2014)
	Advancing the next generation of health risk assessment by Ila Cote, Paul Anastas, Linda Birnbaum, Becki Clark, David Dix, Stephen Edwards, and Peter Preuss (2012)
	Summary Report of Advancing the Next Generation of Risk Assessment Public Dialogue Conference by EPA (2011a)
	Advancing the Next Generation (NexGen) of Risk Assessment: The Prototypes Workshop by EPA (2010)
Tier 3 Prototypes: *Leukemia & Benzene, Lung Injury & Ozone, Liver Cancer & B[a]P/PAHs*	Progress in assessing air pollutant risks from *in vitro* exposures: Matching ozone dose and effect in human airway cells by Gary Hatch, Kelly Duncan, David Diaz-Sanchez, Michael Schmitt, Andrew Ghio, Martha Carraway, John McKee, Lisa Dailey, Jon Berntsen, and Robert Devlin (in press)
	Ozone induces a pro-inflammatory response in primary human bronchial epithelial cells through MAP kinase activation without NF-κB activation by Sean McCullough, Kelly Duncan, Samantha Swanton, Lisa Dailey, David Diaz-Sanchez, and Robert Devlin (in press)
	Characterization of changes in gene expression and biochemical pathways at low levels of benzene exposure by Reuben Thomas, Alan Hubbard, Cliona McHale, Luoping Zhang, Stephen Rappaport, Qing Lan, Nathaniel Rothman, Kathryn Guyton, Roel Vermeulen, Jennifer Jinot, Babasaheb Sonawane, and Martyn Smith (2014)
	Temporal profile of gene expression alterations in primary human bronchial epithelial cells following *in vivo* exposure to ozone by Kelly Duncan, James Crooks, David Miller, Lyle Burgoon, Michael Schmitt, Stephen Edwards, David Diaz-Sanchez, and Robert Devlin (2013)
	IRIS Toxicological Review of Benzo[a]pyrene (Public Comment Draft). U.S. Environmental Protection Agency (2013d), Washington, DC, EPA/635/R-13/138a-b
	Current understanding of the mechanism of benzene-induced leukemia in humans: Implications for risk assessment by Cliona McHale, Luoping Zhang, and Martyn Smith (2012)
	Benzene, the exposome and future investigations of leukemia etiology by Martyn Smith, Luoping Zhang, Cliona McHale, Christine Skibola, and Stephen Rappaport (2011)
	Global gene expression profiling of a population exposed to a range of benzene levels by Cliona McHale, Luoping Zhang, Qing Lan, Roel Vermeulen, Guilan Li, Alan Hubbard, Kristin Porter, Reuben Thomas, Christopher Portier, Min Shen, Stephen Rappaport, Songnian Yin, Martyn Smith, and Nathaniel Rothman (2010)
Tier 2 Prototypes *Knowledge Mining Diabetes/Obesity*	Building associations between markers of environmental stressors and adverse human health impacts using frequent itemset mining by Shannon Bell and Stephen Edwards (2014)
	Systematic identification of interaction effects between genome- and environment-wide associations in type 2 diabetes mellitus by Chirag Patel, Rong Chen, Keiichi Kodama, John Ioannidis, and Atul Butte (2013)
	Data-driven integration of epidemiological and toxicological data to select candidate interacting genes and environmental factors in association with disease by Chirag Patel, Rong Chen, and Atul Butte (2012a)
	Genetic variability in molecular responses to chemical exposure by Chirag Patel and Mark Cullen (2012)
	Role of environmental chemicals in diabetes and obesity: A National Toxicology Program workshop review by Kristina Thayer, Jerrold Heindel, John Bucher, and Michael Gallo (2012)
Tier 2 Prototypes *Short term in Vivo Nonmammalian*	Current perspectives on the use of alternative species in human health and ecological hazard assessments by Edward Perkins, Gerald Ankley, Kevin Crofton, Natàlia Garcia-Reyero, Carlie LaLone, Mark Johnson, Joseph Tietge, and Daniel Villeneuve (2013)
	Propiconazole inhibits steroidogenesis and reproduction in the fathead minnow (*Pimephales promelas*) by Sarah Skolness, Chad Blanksma, Jenna Cavallin, Jessica Churchill, Elizabeth Durhan, Kathleen Jensen, Rodney Johnson, Michael Kahl, Elizabeth Makynen, Daniel Villeneuve, and Gerald Ankley (2013)
	Zebrafish developmental screening of the ToxCast™ Phase I chemical library by Stephanie Padilla, Daniel Corum, Beth Padnos, Deborah Hunter, Andrew Beam, Keith Houck, Nisha Sipes, Nicole Kleinstreuer, Thomas Knudsen, David Dix, and David Reif (2012)
	A systems toxicology approach to elucidate the mechanisms involved in RDX species-specific sensitivity by Christopher Warner, Kurt Gust, Jacob Stanley, Tanwir Habib, Mitchell Wilbanks, Natàlia Garcia-Reyero, and Edward Perkins (2012)

Tier 2 Prototypes ***Short term In Vivo Mammalian***	Incorporating new technologies into toxicity testing and risk assessment: Moving from 21st century vision to a data-driven framework by Russell Thomas, Martin Philbert ,Scott Auerbach, Barbara Wetmore, Michael DeVito, Ila Cote, Craig Rowlands, Maurice Whelan, Sean Hays, Melvin Andersen, Bette Meek, Lawrence Reiter, Jason Lambert, Harvey Clewell III, Martin Stephens, Jay Zhao, Scott Wesselkamper, Lynn Flowers, Edward Carney, Timothy Pastoor, Dan Petersen, Carole Yauk, and Andy Nong (2013c)
	Temporal concordance between apical and transcriptional points of departure for chemical risk assessment by Russell Thomas, Scott Wesselkamper, Nina Wang, Jay Zhao, Dan Peterson, Jason Lambert, Ila Cote, Longlong Yang, Eric Healy, Michael Black, Harvey Clewell III, Bruce Allen, and Melvin Andersen (2013d)
	Integrating pathway-based transcriptomic data into quantitative chemical risk assessment: A five chemical case study by Russell Thomas, Harvey Clewell III, Bruce Allen, Longlong Yang, Eric Healy, and Melvin Andersen (2012c)
	Application of transcriptional benchmark dose values in quantitative cancer and noncancer risk assessment by Russell Thomas, Harvey Clewell III, Bruce Allen, Scott Wesselkamper, Nina Wang, Jason Lambert, Janet Hess-Wilson, Jay Zhao, and Melvin Andersen (2011)
Tier 1 Prototypes ***Integration of QSAR and Various Biological Data Streams***	Developmental toxicity prediction by Raghuraman Venkatapathy and Nina Wang (2013)
	Predictive QSAR modeling: Methods and applications in drug discovery and chemical risk assessment by Alexander Golbraikh, Xiang Simon Wang, Hao Zhu, and Alexander Tropsha (2012)
	Predictive modeling of chemical hazard by integrating numerical descriptors of chemical structures and short-term toxicity assay data by Ivan Rusyn, Alexander Sedykh, Yen Low, Kathryn Guyton, and Alexander Tropsha (2012)
	Application of computational toxicological approaches in human health risk assessment I. A tiered surrogate approach by Nina Wang, Jay Zhao, Scott Wesselkamper, Jason Lambert, Dan Petersen, and Janet Hess-Wilson (2012b)
	An *in silico* approach for evaluating a fraction-based, risk assessment method for total petroleum hydrocarbon mixtures by Nina Wang, Glenn Rice, Linda Teuschler, Joan Colman, and Raymond Yang (2012c)
	Development of quantitative structure-activity relationship (QSAR) models to predict the carcinogenic potency of chemicals. II. Using oral slope factor as a measure of carcinogenic potency by Nina Wang, Raghuraman Venkatapathy, Robert Mark Bruce, and Chandrika Moudgal (2011)
Tier 1 Prototypes ***High throughput Screening***	*In vitro* and modelling approaches to risk assessment from the U.S. Environmental Protection Agency ToxCast programme by Richard Judson, Keith Houck, Matt Martin, Thomas Knudsen, Russell Thomas, Nisha Sipes, Imran Shah, John Wambaugh, and Kevin Crofton (2014)
	Perspectives on validation of high-throughput assays supporting 21st century toxicity testing by Richard Judson, Robert Kavlock, Matthew Martin, David Reif, Keith Houck, Thomas Knudsen, Ann Richard, Raymond Tice, Maurice Whelan, Menghang Xia, Ruili Huang, Christopher Austin, George Daston, Thomas Hartung, John Fowle III, William Wooge, Weida Tong, and David Dix (2013)
	Estimating toxicity-related biological pathway altering doses for high-throughput chemical risk assessment by Richard Judson, Robert Kavlock, Woodrow Setzer, Elaine Cohen Hubal, Matthew Martin, Thomas Knudsen, Keith Houck, Russell Thomas, Barbara Wetmore, and David Dix (2011)
Key Risk Assessment Issues	The role of advanced biological methods and data in regulatory rationality of risk-based regulatory decisions by Douglas Crawford-Brown (2013)
	Incorporating new technologies into toxicity testing and risk assessment: Moving from 21st century vision to a data-driven framework by Russell Thomas, Martin Philbert, Scott Auerbach, Barbara Wetmore, Michael DeVito, Ila Cote, Craig Rowlands, Maurice Whelan, Sean Hays, Melvin Andersen, Bette Meek, Jason Lambert, Harvey Clewell III, Martin Stephens, Jay Zhao, Scott Wesselkamper, Lynn Flowers, Edward Carney, Timothy Pastoor, Dan Petersen, Carol Yauk, and Andy Nong (2013c)
	Addressing human variability in next-generation human health assessments of environmental chemicals by Lauren Zeise, Frederic Bois, Weihsueh Chiu, Dale Hattis, Ivan Rusyn, and Kathryn Guyton (2012)
	Quantitative high-throughput screening for chemical toxicity in a population-based *in vitro* model by Eric Lock, Nour Abdo, Ruili Huang, Menghang Xia, Oksana Kosyk, Shannon O'Shea, Yi-Hui Zhou, Alexander Sedykh, Alexander Tropsha, Christopher Austin, Raymond Tice, Fred Wright, and Ivan Rusyn (2012)
	Predicting later-life outcomes of early-life exposures by Kim Boekelheide, Bruce Blumberg, Robert Chapin, Ila Cote, Joseph Graziano, Amanda Janesick, Robert Lane, Karen Lillycrop, Leslie Myatt, Christopher States, Kristina Thayer, Michael Waalkes, and John Rogers (2012)
	In vitro screening for population variability in chemical toxicity by Shannon O'Shea, John Schwarz, Oksana Kosyk, Pamela Ross, Min Jin Ha, Fred Wright, and Ivan Rusyn (2011)

Executive Summary

Over the past 40 years, the U.S. Environmental Protection Agency (EPA) has made significant progress in protecting human health and the environment from the adverse effects of chemical exposures. The tens of thousands of chemicals in the environment, however, are overwhelming our ability to evaluate their safety using traditional approaches. Traditional methods also are not adequately addressing complex, risk assessment issues such as co-exposures from many different environmental stressors or the potential effects of chemicals on people who might be more sensitive or susceptible. This report, *Next Generation Risk Assessment: Recent Advances in Molecular, Computational, and Systems Biology* (NexGen), explores new, more efficient approaches to evaluating chemical safety and to addressing key issues. Applications range from screening and prioritizing thousands of chemicals for further evaluation to augmenting traditional, data-rich chemical assessments in support of national regulations. This report presents the results from a multiyear, multi-organization effort designed to summarize the state of the science and to provide a scientific foundation for modernizing risk assessment methodology. The target audience for this report is scientists and risk assessors already familiar with many of the technical terms, concepts, and practices discussed. This executive summary, however, offers a less technical overview of the contents of this report. Additional general information is available at the NexGen website (EPA 2013a).

Eight case studies, or prototypes, were developed to illustrate how new science might be used to support a variety of Agency decisions. Areas of interest include the following: Can new types of data produce results comparable to the results of traditional risk assessments? What types of information appear most valuable for specific purposes? And what are the decision rules needed during the selection and evaluation of new data types to ensure consistent, scientifically sound assessments? The prototypes are not intended to evaluate all the available new data and methods or all situations risk managers face. Rather, the intent is to provide concrete examples of some analyses and uses, to encourage further dialogue, and to promote broader understanding of these new risk assessment approaches.

Based on the lessons learned from developing the prototypes, several near-term and long-term implications for risk assessment can be highlighted as follows:

1. Significant progress has been made in implementing the vision presented in the National Research Council's (NRC) report, *Toxicity Testing in the 21st Century*, and in EPA's report, *Strategic Plan for the Future of Toxicity Testing and Risk Assessment at the U.S. Environmental Protection Agency*.

2. New tools and data types are facilitating, on an unprecedented scale, the testing and evaluation of chemicals that previously could not be evaluated due to limited or no traditional toxicity data.

3. New data are dramatically improving our understanding of the causes of disease, the effects from low levels of exposures, and why certain people might be more susceptible to chemical effects (e.g., due to differences in age, health status, or genetics).

4. These new data types are being organized, curated, stored, and made publicly available in massive data warehouses.

5. Advanced, automated approaches are being developed to analyze these large data sets rapidly and comprehensively.

6. This new knowledge provides powerful insights into the contribution of environmental risk factors to adverse health outcomes and can be used to better inform risk assessment.

Substantial scientific challenges and uncertainties, however, remain. Specific challenges include: (1) developing reliable, predictive molecular indicators or biomarkers of exposure and effects for a wide variety of chemicals; (2) understanding the key pathways in the network of interactions among genes, cells, tissues, and organs that is needed to conduct predictive toxicology; (3) further characterizing human variability and how genetic makeup, preexisting backgrounds of disease and exposure, and adaptive or compensatory processes combine to influence population risks; (4) accounting for variables in test systems that can influence observed associations between molecular perturbations and disease outcomes (e.g., experimental design, metabolism, genomic variants, target cell type(s), cell and tissue interactions, species, and lifestage); (5) understanding the role of epigenomics in risks; and (6) characterizing, in the best way possible, the uncertainties and confidence in risk assessments informed by new data types.

As the above challenges are addressed, we anticipate that the new approaches discussed in this report will provide a variety of applications to risk managers within EPA, and the risk assessment community at large, including identifying safer chemicals and processes, reducing hazardous chemicals in the environment, and improving our ability to protect public health and the environment. The scientific community and the public should anticipate transitions to new types of risk assessments over several years, particularly for screening and prioritization of large numbers of chemicals and support of nonregulatory decision-making. A variety of new tools with various associated uncertainties will be evaluated in differing applications, externally reviewed, and refined. Near-term progress will include case-by-case development of additional examples for peer and public review and workshops to help inform critical issues. EPA's Chemical Safety for Sustainability (CSS) and Human Health Risk Assessment (HHRA) research program plans and the National Institute of Environmental Health Sciences' (NIEHS) Strategic Plan address many of the research implications discussed in this report.

Careful evidence integration will continue to be required for NexGen-informed assessments, as has been the case for traditional assessments. Traditional approaches for systematic review and evaluation of evidence are being adapted and applied to the new types of data, to ensure data quality, transparency, and confidence in the overall evidence. The hurdles to providing convincing evidence that a chemical causes or contributes to an adverse outcome, however, are substantial. Thus, for the foreseeable future, major risk assessments used to support national regulation will continue to be based on traditional data, although increasingly augmented by new data as confidence increases in the predictive capability of these new approaches.

Lastly, significant outreach, education and interaction with our stakeholders will continue to be a priority for EPA to ensure the transparency of new science, and to improve our understanding of how best to apply these advances to environmental health risk assessment.

1 Introduction

Over the past 40 years, the U.S. Environmental Protection Agency (EPA) has made significant progress in protecting public health and the environment from the adverse effects of chemical exposures. The tens of thousands of chemicals in the environment, however, are overwhelming our ability to evaluate their safety using traditional approaches. Traditional methods also are not adequately addressing complex, risk assessment issues such as co-exposures from many different environmental stressors, or the potential effects of chemicals on people who might be more sensitive or susceptible. This report, *Next Generation Risk Assessment: Recent Advances in Molecular, Computational, and Systems Biology* (NexGen), explores new approaches that are faster and less resource intensive than traditional approaches and hold great promise in addressing these problems. This report is the culmination of a multiyear, multi-organization effort involving five U.S. federal agencies and three European agencies, Health Canada, California Environmental Protection Agency, Hamner Institutes for Health Sciences, scientists from 12 universities, and several other organizations that provided staff, data, advice, and review.[1,2] Specific aims for the NexGen effort are noted in Box 1.

Box 1. Specific Aims of NexGen

- Consider how new risk assessment approaches might inform particular risk management situations (decision context) to create "fit for purpose" assessments.
- Develop prototypes that illustrate uses of new data types and methods to better inform risk assessment.
- Understand what data types are most informative for a given situation (value of information).
- Adapt existing decision rules for use with new data types and approaches, thus ensuring consistent, scientifically defensible assessments
- Identify issues, challenges, and next steps.

Recent scientific and technological advances are providing unprecedented opportunities to understand human environmental risks. Massive amounts of new data are being generated, often using robotics (Derry et al. 2012; Friend 2013; Sturla et al. 2014).[3] These data are stored, managed, curated and made publicly available in several data warehouses, such as those in the National Institutes of Health National Library of Medicine (Chadwick 2012; Collins 2009; Kleinberg and Hripcsak 2011; Mechanic et al. 2012; NCBI 2014b, d) and the EPA Aggregated Computational Toxicology Resource (Dix et al. 2007; EPA 2014d, l; Judson et al. 2012). Concomitantly, powerful new bioinformatic methods are being developed to identify, organize, and analyze these data. Profound insights are beginning to emerge into the causes of disease, the contributions of environmental factors, and what might make individuals and subpopulations

[1]Appendix A summarizes ongoing work at several government agencies to advance the next generation of toxicity testing and risk assessment.
[2]The government participants in this effort also are working through the Office of Economic Cooperation and Development and the World Health Organization (WHO) to redesign toxicity testing and risk assessment of chemicals in the environment and to harmonize approaches worldwide (EC 2013; JRC 2014; Meek et al. 2014; NIEHS 2014a; OECD 2010, 2014d; Sturla et al. 2014; Thomas, R. S. et al. 2013b; Tice et al. 2013).
[3]Approximately 1.8 zettabytes (10^{21}) of new data from tens of thousands of new papers are generated every year, roughly doubling the world's information every two years (Dearry 2013).

susceptible (Bhattacharya et al. 2011; Chiu et al. 2010). Two examples of new types of data collection, integration, and interpretation are:

- Tox21/ToxCast™,[4] which is developing new assays for chemical safety and testing 10,000 chemicals (Figure 1) (Attene-Ramos et al. 2013; EPA 2014l; Jacobs 2011; Judson et al. 2014; Tice et al. 2013).
- The continuing characterization of genomes, epigenomes, and environment-wide associations with disease in tens of thousands of humans (ENCODE Project Consortium 2012; Friend 2013; Mechanic et al. 2012; The 1000 Genomes Project Consortium 2010).

Figure 1. Toxicology Testing in the 21st Century (Tox21) Robot Conducts Bioassays on 10,000 Chemicals.
A robot arm (foreground) retrieves assay plates from incubators and places them at compound transfer stations or hands them off to another robot arm (background) that services liquid dispensers or plate readers. Photo by Maggie Bartlett (NHGRI 2014b).

Such large-scale knowledge creation was unimaginable 15 years ago.

This report is not an exhaustive survey of all data relevant to the prototypes or of all new approaches in this rapidly developing area of science. Rather, it highlights some of the most interesting and promising approaches and identifies challenges to their use in risk assessment. Forty papers and reports were developed specifically for this effort (see list on pages xiv-xvi) and more than 450 references provide additional scientific technical details.

This effort represents an important step in the implementation of the National Research Council's (NRC) *Toxicity Testing in the 21st Century* and *Science and Decisions: Advancing Risk Assessment*, and EPA's *Strategic Plan for the Future of Toxicity Testing and Risk Assessment at the U.S. Environmental Protection Agency*. Importantly, the NexGen provides a scientific basis for modernizing risk assessment. Responses to both peer-review comments and public comments on the September 2013 draft report are incorporated in this final report.

This NexGen program report is organized as follows:

- Section 1: Introduction.
- Section 2: Preparation for Prototype Development – describes preliminary work, including planning for "fit-for-purpose" assessments (decision context); reports on an overarching framework, describes interactions with experts and stakeholders, and develops key questions to be addressed, and considers systematic review and evidence integration.

[4]Tox21 stands for the Toxicology in the 21st Century program, and ToxCast stands for Toxicity Forecaster.

- Section 3: The Prototypes – presents detailed examples of using various advanced methods and data to consider the questions developed in Section 2. The prototypes are matched to three categories of decision contexts (also discussed in Section 2), starting with the *in vitro* and *in vivo* data-rich chemicals (Tier 3), proceeding to chemicals with robust *in vitro* and limited *in vivo* data (Tier 2), followed by chemicals that have robust *in vitro* data but very limited or no *in vivo* data to support traditional risk assessment (Tier 1).
- Section 4: Advanced Approaches to Recurring Issues in Risk Assessment – discusses how advanced methods can be used to address ongoing challenging issues, such as human population variability and sensitivity, cumulative risk, and responses at environmental levels.
- Section 5: Lessons Learned from Developing the Prototypes – reviews and summarizes what the prototype development process taught us.
- Section 6: Challenges and Research Directions – looks to challenges that must be met to further new testing and risk assessment, and planned research.
- Section 7: References
- Appendix A summarizes ongoing activities at several government agencies in the United States and Europe that are providing additional data and analyses related to advancing NexGen.
- Appendix B provides details of interactions with the scientific community and stakeholders. Appendix C lists recommended principles and methods for uncertainty and variability analysis.
- Appendix D provides a glossary of terms used throughout this report.

2 Preparation for Prototype Development

2.1 Planning for Fit-for-purpose Assessments

EPA needs various types of risk assessments to address different situations or decision contexts.[5] We designed the prototypes around broad categories of potential end uses and with the intent of developing "fit-for-purpose" assessments. Fit-for-purpose simply means that a product meets the needs of the end user. The categories we chose greatly oversimplify the types of decisions risk managers face, but hopefully they will illustrate how new approaches could be used. In reality, these approaches represent a set of tools that can be used to support a variety of decisions (EPA 2014i). The illustrative categories used in this report are:

- **major-scope decision-making** – generally regulatory decisions;
- **limited-scope decision-making** – usually nonregulatory decisions; and
- **prioritization and screening decisions** – ranking chemicals for additional evaluation, and urgent response.

These categories reflect a range of environmental challenges—from the need to screen many untested chemicals in the environment to the need to implement national regulations for high-profile chemicals. Figure 2 presents characteristics of the three decision categories and examples of potential prototype applications. These decision context categories were developed during discussions among EPA risk assessors and managers (EPA 2011b). Three factors integral to the decision context for risk managers are the (1) magnitude and prevalence of potential exposures, (2) numbers of chemicals to be considered, and (3) weight of scientific evidence required for specific types of decision-making. Both legislative mandates and historical precedence are important influences on the decision context and specific regulatory actions.

Three examples of previous decisions that used new data types, and that could be considered major scope, limited scope, and prioritization and screening are the (1) International Agency for Research on Cancer's determination on a likely causal link between benzene exposures and lymphoma based on molecular mechanisms data (IARC 2012); (2) examination of cumulative risk potential from relatively uncharacterized conazole fungicides based on molecular mechanisms data (EPA 2011d; Hester et al. 2011); and (3) Deep Water Horizon/Gulf of Mexico oil spill dispersants using *in vitro* high-throughput data (Judson et al. 2010).

[5] "**Decision context**" is defined as the circumstances that form the setting for decision-making, and in terms of which the decision can be assessed and understood. More simply put, to characterize the decision context, one asks the questions, "What decision options are being considered?" and "What products or information are needed to support those decisions?" (NRC 2009).

Example Decision Context Categories for Which Illustrative "Fit for Purpose NexGen Assessment Prototypes Were Developed			
	Tier 1 Prioritization and Screening	**Tier 2 Limited scope Decision making**	**Tier 3 Major scope Decision making**
Common Characteristics	• Exposures assumed due to use in commerce • Very limited or no traditional hazard data • 10,000s of chemicals of interest	• Some specific inventory of chemicals, monitored or modeled exposure data • Potentially some limited traditional data • 1000s of chemicals of interest	• Generally widespread, demonstrated exposures • Extensive traditional data; unresolved issues could remain • 100s of chemicals of interest
Possible Applications	• Situations where large numbers of chemicals require sorting for further action • Lifecycles, sustainable chemical and process evaluations • Emerging issues evaluation • New assessment queuing • Urgent or emergency response • Research or testing priority setting	• Superfund remediation/ hazardous waste disposal • Water contaminants identification • Urban air toxins assessment • Chemical mixture evaluations • New assessment queuing • Urgent or emergency response • Research or testing priority setting	• High-profile, nationally important assessments • Community assessments • Research or testing priority setting
Increasing: exposure potential, weight of scientific evidence required for decision-making, resources required for assessment. Decreasing numbers of chemicals evaluated or assessed.			

Figure 2. Description of General Decision Context Categories Suggested by EPA Program Offices.

2.2 A Framework

The second task in planning the NexGen prototypes was to develop an assessment framework. The NexGen framework incorporated several essential elements of earlier risk assessment frameworks and provided guiding principles for the NexGen effort. A draft version of this framework was presented and discussed in a November 2010 meeting with scientific experts (EPA 2010) and again in a February 2011 public meeting with stakeholders (EPA 2011a). Feedback from scientific experts and the public helped refine the framework. The final version represents a continued evolution and is described in detail in Krewski et al. (2014). This section is adapted from Krewski et al. (2014).

Key elements of risk science and population health are combined in the NexGen framework to provide a multidisciplinary approach to assessing and managing health risk issues (Krewski et al. 2007). The framework presented in Figure 3 is built on three cornerstones: (1) new risk assessment methodologies that consider new data types and inform risk management decision-making; (2) new data types from advances in molecular, computational, and systems biology aimed at understanding perturbations in biological pathways that lead to adverse effects; and (3) a population health perspective that recognizes that most adverse health outcomes involve multiple determinants (i.e., multiple causal or contributing factors).

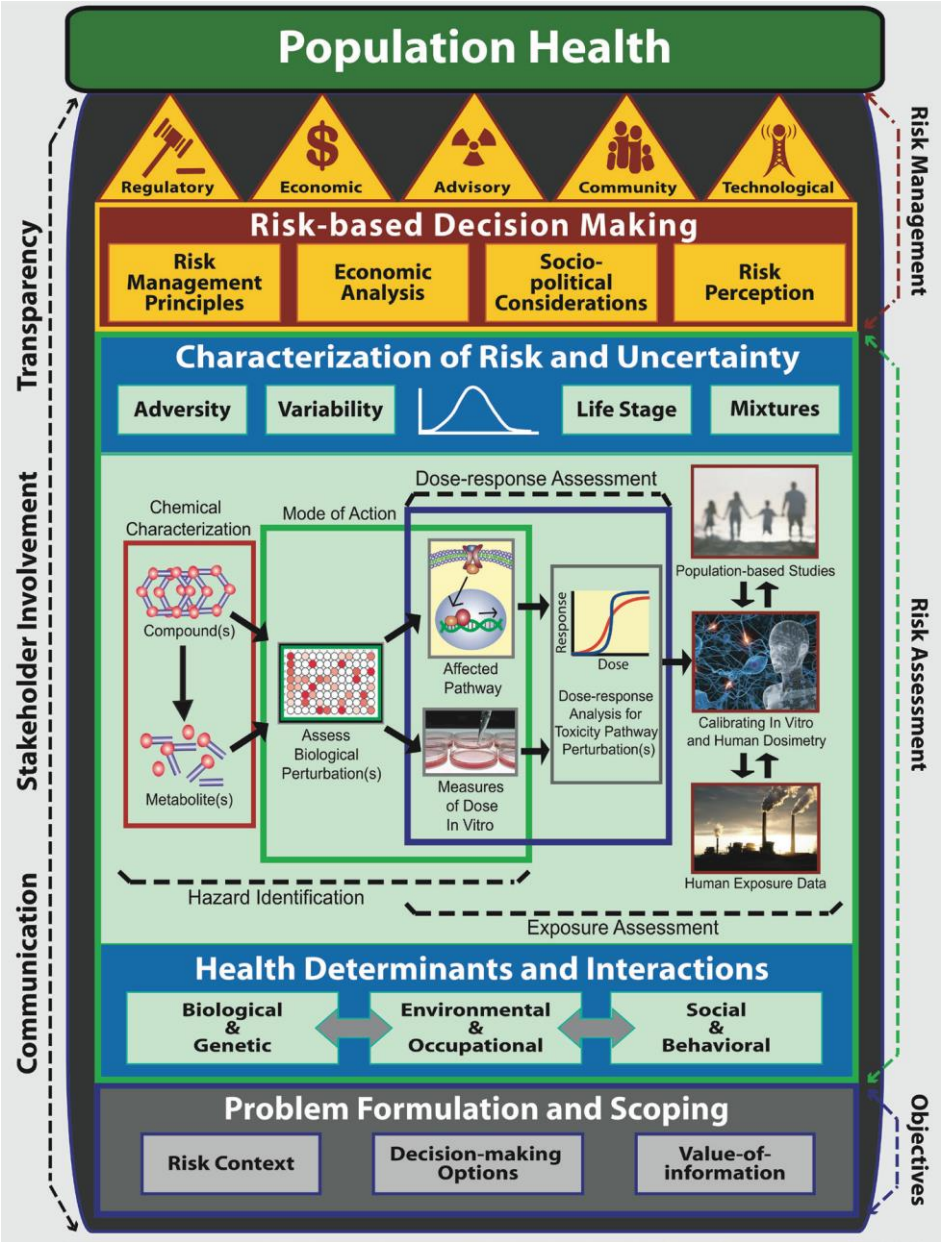


Figure 3. The Next Generation Framework for Risk Science.
Phase I: objectives—problem formulation and scoping takes into consideration the risk context, decision-making options, and value of information. Phase II: risk assessment: health determinants and interactions—incorporates a population health approach that takes into account multiple health determinants that interact with the risk factor(s) of interest. Hazard identification, dose–response assessment, and exposure assessment make use of new scientific tools and technologies, based on high throughput screening assays and computational methods in biology and toxicology for hazard identification and dose–response assessment; in vitro to in vivo extrapolation methods for calibration of in vitro and human dosimetry; molecular and genetic epidemiology to identify toxicity pathway perturbations in population-based studies; and high-performance mass spectrometry to generate human exposure data, to assess risk characterization of risk and uncertainty applies new risk assessment methodologies to develop human exposure guidelines. Phase III: risk management—risk-based decision-making considers fundamental risk management principles, economic analysis, sociopolitical consideration and risk perception to select one or more risk management interventions of a regulatory, economic, advisory, community-based, or technological nature for risk management. (The center section on hazard identification, dose–response assessment, and exposure assessment is adapted from Figure 2 of Krewski et al. 2011.)

2.3 Science Community and Stakeholder Engagement

The third task was to reach out to the science community and stakeholder groups to communicate our plans and to benefit from their input. Outreach is an essential principle of the framework described in Section 2.2. Our outreach involved many efforts. We (1) convened an experts workshop to review the prototype concepts (2010); (2) sponsored a public dialogue conference to communicate our plans and elicit feedback from the public (see Figure 4)(2011); (3) evaluated and incorporated results from academic surveys of the business community and the environmental communities (2011 and 2012); (4) hosted a NexGen website to communicate activities and progress (EPA 2013a); (5) participated in a National Academy of Sciences – Emerging Science workshop (2012); (6) participated in Advisory Board/Board of Scientific Counselors meetings (2012, 2014); and (7) elicited and responded to external peer-review and public comment on the draft document. (See Appendix B for more details on interactions with the scientific community and stakeholders.)

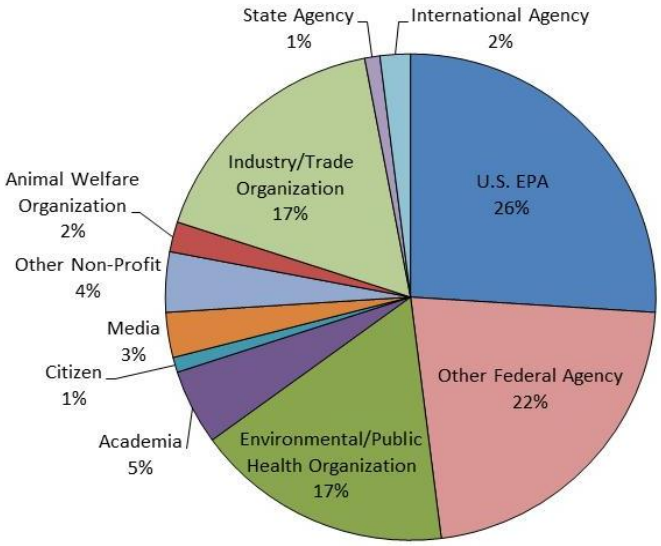


Figure 4. Categories of Stakeholders that Attended the February 2011 NexGen Public Dialogue Conference (EPA 2011a)

Comments from the scientific community and stakeholders on advancing new methods in risk assessment were generally positive, although substantial and various concerns were expressed. Experts in molecular, computational, and systems biology were generally very optimistic that new data types could inform risk assessment. The public-interest groups and the business community recognized the potential to evaluate chemicals more efficiently, and were guardedly optimistic, yet had concerns about the specifics of application and interpretation. Their concerns included an interest in demonstrations on the value of new approaches, caution about the potential to overstate the utility and efficiency of NexGen approaches, questions about how NexGen prototypes will address key methodological issues, the need for transparency and meaningful public engagement; how the results would be used in risk management, and if timely and effective communications would occur. Some in the business community expressed concern over whether EPA could develop the necessary expertise to guide the program to a successful conclusion. EPA, experts, and stakeholder groups all recognize the challenges ahead and the need for continued interactions. Winning over a larger community less familiar with the complex science associated with new approaches, and potentially more skeptical, will likely be challenging, but such challenges are considered surmountable if EPA can build capacity and communicate effectively how the approaches can be used in risk assessment.

2.4 Recurring Issues in Risk Assessment

The fourth task that preceded the actual prototype development was to identify recurring issues that new methods and data might substantively inform. The issues identified included problem formulation, evidence integration for hazard (formerly termed weight-of-evidence[6]) and dose-response (including internal dosimetry) estimation, characterizing variability in human response, interspecies extrapolation, cumulative risk assessment, and uncertainty characterization. The prototypes were evaluated in this context, and Section 4 discusses the insights gained from this exercise.

2.5 Key Questions and Evidence Integration

Through the efforts described above, a set of questions (Table 1) was developed to guide the prototype development and evaluation.

For an activity as critical as risk assessment, studies selected for consideration should be well designed, carefully conducted, and transparently reported, in accordance with traditional practices (EPA 2005, 2013c; NRC 2014; U.S. DHHS 2014). Systematic review of available data and evidence integration was considered in advance of the prototype development. Evidence from selected studies was integrated and used to evaluate causality. The evidence for causality is increased by consistency of the data across multiple, independent studies, the coherence of the data across different data types, and the biological plausibility of the association between cause and effect. Chance, bias, and confounding should be ruled out or minimized with reasonable confidence to infer a causal or likely causal relationship. When chance, bias, or confounding cannot be minimized, data are "suggestive" or "insufficient." Adaptations of the Bradford-Hill "criteria" continue to prove useful in evaluating data (EPA 2005, 2013e; Hill 1965; Meek et al. 2014; U.S. DHHS 2014). Kleinberg and Hripcsak (2011) provide additional discussion on systematic review and evidence integration as it specifically applies to new data types. Examples of the types of evaluations that could provide sufficient evidence to infer causal or likely causal relationships among exposure, molecular events, and adverse outcome include:

- meta-analyses of multiple well-conducted studies that provide consistent findings of significant associations among exposures, intermediate effects, and outcomes; and
- experimental studies that identify pollutant-induced modification in specific pathways or networks coupled showing that these modifications alter adverse outcomes, for example,

[6]In the recent NRC review (2014) of the EPA's Integrated Risk Assessment System (IRIS) process "the committee found that the phrase *weight of evidence* has become far too vague as used in practice today and thus is of little scientific use…The present committee found the phrase *evidence integration* to be more useful and more descriptive of what is done" in EPA's major assessments….that is, "assessments must come to a judgment about whether a chemical is hazardous to human health and must do so by integrating a variety of evidence (see Figure 6-1 of the NRC report for more details).

pharmacological interventions that block exposure-dependent pathway alterations and, concomitantly, block or mitigate adverse outcomes;

- traditional data (e.g., whole animal bioassay data) augmented by molecular biology, such as mechanistic information;
- identification of idiopathic gene variants that alter the risks of adverse outcomes and provide evidence linking pathways to outcomes; and (Q)SAR comparisons of the molecular data from sufficiently similar chemicals to infer associations among exposures, molecular pathway alterations and adverse outcomes.

Table 1. Questions Posed in Regard to the Prototypes

Tier	Hazard Identification Questions	Exposure Dose Response Questions	Potential Applications
3	• How can adverse outcome pathway (AOP) networks be used to characterize environmentally related human disease or disorder? • Can AOP networks also be used to identify chemicals and nonchemical stressors that operate by the same mechanism and, thus, should be considered together? • Can gene variants be identified that are hallmarks for susceptible subpopulations? • How can this information be extended to the evaluation of relatively unstudied chemicals?	• Can an AOP network or components of a network be used as biomarker of exposure or dose and/or effect? • Can AOP networks be used to characterize the combined risks from chemicals or nonchemical stressors? • Can differential sensitivity of subpopulation to chemical exposures be characterized? • How can this information be extended to the evaluation of relatively unstudied chemicals?	• To use AOPs developed from human data to screen relatively unstudied chemicals. • To use human-derived AOPs to verify AOPs developed in nonhumans, short-duration *in vivo* or *in vitro* exposure studies • To address key traditional unresolved data gaps, such as low exposure-dose response or species to species extrapolation • To increase the evidence for cause and effect through mechanistic knowledge
2	• Can knowledge mining or short-term *in vivo* approaches efficiently identify potential hazard? • Can these new medium-throughput approaches help describe AOPs or AOP networks? • How could this medium-throughput based information be used in risk assessment?	• Can potency be reliably estimated for human risks? • Is the estimated toxicity value an absolute or relative potency? • What models, methods, and data are needed to estimate human equivalent dose?	• To screen hundreds to thousands of relatively unstudied chemicals for hazard and relative or absolute potency • To maximize use of very large existing data sets (potentially all published data) • To screen for cumulative risk potential
1	• How can *in vitro* approaches be used effectively to screen many thousands of chemicals for potential hazard? • Can new high throughput approaches help identify AOPs or AOP networks? • How can this high-throughput based AOP information be used in risk assessment?	• What endpoints can be used reliably to evaluate potential toxicity? • Is the toxicity value absolute or relative potency? • What is needed to estimate human equivalent dose? • What is needed to extrapolate to human population risks?	• To screen thousands to tens of thousands of relatively unstudied chemicals for hazard and relative or absolute potency • To maximize use of very large existing data sets (potentially all published data) • To screen for cumulative risk potential

Lastly, due to the complexities of biology, linking disruption of normal biological processes to a specific disease or disease risk is challenging. Ranking chemicals based on their potency to alter biological processes, however, appears possible without knowing how or if such disruptions will be reflected in terms of disease risks. Characterizing potency, without the clear identification of

hazard, is a reversal of the traditional risk assessment approach (i.e., hazard identification followed by dose-response assessment). High- and medium- throughput methods are being developed to evaluate chemical potencies in this way, particularly for sorting large numbers of chemicals based on potential concern. The Tox21 and ToxCast programs are examples of such efforts. As our mechanistic understanding of the links between molecular events and human disease and disorder improves, these data and their predictive capability will become increasingly useful.

3 The Prototypes

In their *21st Century* toxicity testing and risk assessment strategy documents, both the National Academy of Sciences and EPA recommended prototype development and identified key issues to consider (EPA 2009b; NRC 2007a, 2009). These recommendations were used as a starting point for the NexGen program. The scope of the prototypes and key questions considered were developed from discussions with EPA Program Offices, the partner organizations, and science experts; these discussions led to the prototype selection criteria shown in Box 2.

Box 2. Selection Criteria for Prototypes

- Decision context applicability (i.e., illustrative of "fit for purpose" assessments).
- Multiple, high quality molecular biology studies available.
- Robust traditional data available to compare with conclusions drawn from NexGen data.
- Overall, consistent, coherent, and biologically plausible data available.
- Active collaborations with investigators to benefit from their knowledge, ability to execute additional experiments and analyses as needed.
- Cross organizational and sectors of the risk assessment community collaborations fostered.

Eight case studies or prototypes were developed to illustrate potential uses of new science in supporting a variety of Agency decisions. The prototypes explored the following: whether new types of data can produce results comparable to the results of traditional risk assessments; what types of information appear most valuable for specific purposes; what decision rules are needed when selecting and evaluating new data types to ensure consistent, scientifically sound assessments; and what the challenges are to interpreting and using new data in risk assessment. The prototypes do not consider all data and methods, or situations faced by risk managers. Rather, the intent is to provide illustrative, concrete examples of analyses that encourage further dialogue and that advance our understanding of new risk assessment data and methods. The eight prototype assessments developed for this report,[7] categorized by decision context, are:

- Tier 3: major-scope decision-making prototypes that developed proof of concept and explored augmentation of very traditional data-rich chemical assessments
    - Hematotoxicity and leukemia: benzene and other leukemogens
    - Lung inflammation and injury: ozone
    - Lung and liver cancer: benzo[a]pyrene (B[a]P)/polycyclic aromatic hydrocarbons (PAHs)
- Tier 2: limited-scope decision-making prototypes that explored approaches to assessing hundreds to a few thousand chemicals
    - Diabetes and obesity: knowledge mining and meta-analyses of published literature
    - Thyroid disruption: short-duration, *in vivo* assays—alternative species
    - Cancer- and noncancer-related effects: short-duration, *in vivo* assays—rodent

[7]Different groups, selected for their expertise, developed the various prototypes; consequently, the presentation styles differ somewhat.

- Tier 1: prioritization and screening prototypes that explored approaches to assessing thousands to tens of thousands of chemicals
    - Various environmental contaminants: quantitative structure activity relationship (QSAR) models
    - Various environmental contaminants: high-throughput and high-content *in vitro* assays.

Understanding mechanisms of action[8] in a systems biology context is considered important to understanding new information and fostering new risk assessment applications (Califano et al. 2012; Edwards and Preston 2008; Ideker and Krogan 2012; Mitra et al. 2013; Molinelli et al. 2013; Sturla et al. 2014). To the extent possible, the prototypes were organized around putative mechanisms of disease or disorder. In toxicology, simplified mechanistic models are often termed either modes of action (MOAs) or adverse outcome pathways (AOPs). Models that are somewhat more complex are often termed AOP networks to convey the interconnectedness of AOPs that generally underlie disease. The term AOP sometimes erroneously conveys that toxicity results from novel events rather than perturbations of normal biology. To date, the terminology to discuss mechanistic concepts is not uniform. In particular, the fields of medicine and toxicology use different terminology for similar concepts (e.g., BioSystems versus AOP networks). For consistency, the term AOP network is used throughput this report except in discussions of published works that use other terms. A substantial effort in toxicology is underway to unify descriptions of mechanisms in the context of AOP networks (see Box 3).

Box 3. International Coordination of AOP and MOA Development

Under the auspices of the Organization for Economic Cooperation and Development (OECD), an international program began in 2012 to develop, review, agree on, publish, and endorse adverse outcome pathway constructs. EPA and the European Commission Joint Research Center (JRC) jointly lead this effort, coordinating with the World Health Organization (WHO) s International Programme on Chemical Safety. The effort will foster international consistency, quality, and acceptability of AOPs used for chemical risk assessment (OECD 2014d). While some discussion continues in the science community about potential differences between AOP and MOA, WHO and OECD consider the terms interchangeable (Meek et al. 2014). Nonetheless, WHO, OECD, JRC and EPA have come to an agreement to move toward AOP and AOP network as the preferred terminology.

[8]**Mechanism of action, mode of action (MOA), adverse outcome pathway (AOP) and AOP network** are defined as follows**:** (1) mechanism of action is the complete sequence of biological events that must occur to produce an adverse effect; (2) MOA is defined as a "sequence of key events and processes, starting with interaction of an agent with a cell, proceeding through operational and anatomical changes, and resulting in an adverse health effect"; and (3) AOP describes a "sequential chain of causally linked events at different levels of biological organization that lead to an adverse health or ecotoxicological effect" (OECD 2013, 2014d); and 4) an AOP network is the interrelated AOPs that represent the combination of events and pathways that underlie disease or disorder The term MOA has been in widespread use for several years and was used extensively in the 2009 EPA Cancer Guidelines. In 2012, the Organization for Economic Cooperation and Development launched a new program on the development of AOPs. AOP was chosen as a new term to emphasize use in population risk assessment (Ankley et al. 2010).

Two basic approaches are used to develop systems level understanding: bottom up and top down. The bottom-up approach focuses on molecular and cellular components, and seeks to understand how these components are networked, and how normal network function is altered following exposure to chemicals or stressors. The bottom-up approach generally uses information from new types of *in vitro* testing and some *in vivo* alternative animal testing. This information is used to predict how perturbations at the molecular and cellular levels might propagate. The bottom-up approach is addressed most extensively in Tiers 1 and 2 for chemicals having little or no traditional *in vivo* data, and takes advantage of new, large data sets, such as ToxCast and Tox21. The top-down approach focuses on network interactions and disease indicators at the whole-body or population level, based often on human clinical and epidemiological data, and associations between disease states and environmental factors (Friend 2013). This information is used to identify associated factors at the organ, cell, or molecular level with the potential for a causal relationship with the disease state. This approach is addressed most extensively in Tiers 2 and 3, and often takes advantage of human "big data" sets developed by the National Institutes of Health (NIH) and others, such as BioSystems and the 1000 Genomes Project. Both the bottom-up and top-down approaches are informative, and are best used together to develop integrated and comprehensive knowledge.

A broad array of methods was evaluated in the NexGen prototypes. Tools and techniques used are summarized in Table 2 (Krewski et al. 2014). The assignments of particular methods to decision-context categories in Table 2 are neither fixed nor exclusive. For example, high-content screening (HCS)[9] assays are used primarily in the Tier 2 examples, but they also might be used in Tier 1 screening or in major-scope assessments. As noted earlier, the numbers of chemicals that require evaluation and the decision context are the main considerations in determining the appropriate data and methods to use to design fit-for-purpose assessments (NRC 2009).

An important element of this report is the discussion of the promises and limitations of various approaches being considered and the lessons learned during prototype development.

3.1 Tier 3: Major-scope Assessments

The Tier 3 prototypes focused on chemicals with known public health effects at environmental exposure levels (EPA 2013c, d; IARC 2012). Of particular interest was the examination of causal evidence linking molecular-level data from epidemiological, clinical, or *in vivo* animal exposure studies to the results from traditional *in vivo* assays. The purpose was to test the hypothesis that AOP networks (1) can be identified that are strongly associated with the adverse effects known to result from exposure to the chemicals under study, (2) are exposure-dose dependent within the range of environmental exposures, and (3) can be shown to vary with risk factors such as genomic variants, mixture, and nonchemical stressor exposures. If true, AOP networks could improve our ability to characterize hazards, dose-response, and risk potentially posed by data-limited chemicals,

[9]A **high-content screening (HCS) assay** is defined as any method with multiple simultaneous readouts used to analyze system dynamics at any specified level of organization, but generally referring to the whole body, whole cell, or subcellular level of organization.

Table 2. Prototype Use of New Scientific Tools and Techniques (adapted from Krewski et al. 2014)

Decision Context Category	Tier 1: Prioritization & Screening	Tier 2: Limited scope Assessments	Tier 3: Major scope Assessments
Hazard Identification and Dose response Assessment Methods			
Quantitative structure activity models	■	■	
Toxicity pathways analysis	■	■	■
High throughput *in vitro* assays	■	■	■
High content omics assays		■	■
Biomarkers of effect		■	■
Molecular and genetic population based studies			■
Dosimetry and Exposure Assessment Methods			
In vitro to *in vivo* extrapolation	■	■	
Pharmacokinetic models and dosimetry	■	■	■
Biomarkers of exposure		■	■
Cross cutting Assessment Methods			
Adverse outcome pathways	■	■	■
Bioinformatics and computational biology	■	■	■
Systems biology	■	■	■
Functional genomics		■	■

as well as provide new insights into many historically challenging risk assessment issues, such as identifying human susceptibility and estimating cumulative risks. The most robust data sets identified for this proof of concept exercise were benzene and hematotoxicity/leukemia, ozone and inflammation/lung injury, and tobacco smoke/PAHs/BaP and lung cancer.

Table 3 summarizes the approaches used in the Tier 3 prototypes, and some of the advantages and disadvantages of each approach. Data from a variety of new technologies were evaluated, including deoxyribonucleic acid (DNA) transcription (transcriptomics), protein expression (proteomics), and genome-wide analyses of susceptibility genes (genomic analyses of human gene variants). Bioinformatic analyses (computer-assisted data identification, organization, and synthesis) were used to identify AOP networks and to interpret the molecular data in the context of adverse health outcomes and disease from traditional studies. This integration of new and traditional data provided a relatively detailed picture of causal events from molecular initiation events (MIEs) to intermediate biochemical events to adverse outcomes. Implications for risk assessment identified by the Tier 3 prototypes are discussed at the end of this section and are integrated with other lessons learned in Section 5. Due to the uncertainties associated with new approaches, we

anticipate that major regulatory risk assessment will be based primarily on traditional data for the foreseeable future, albeit augmented by new data types.

Table 3. Summary of Tier 3 NexGen Prototype Approaches, Including Strengths and Weaknesses

TIER 3: MAJOR SCOPE ASSESSMENT PROTOTYPES		
	Benzene and Ozone	**Tobacco Smoke, BaP/PAHs**
Approaches:	• Measurements of chemically induced, dose-dependent alterations in transcriptomics in humans, using specific and sensitive assays • Comparison of molecular epidemiological and clinical studies with concomitantly collected well-characterized adverse health effects • Transcriptomic alterations occurring in genes and pathways correlated with traditional upstream events and adverse effects in same individuals • Well-described human exposures at environmentally relevant concentrations • Measurements of exposure-dose relationships using urinary biomarkers or $^{18}O_2$ dosimetry • Adverse effects can be blocked, partially ameliorated by alterations of implicated genes and pathways • Variability of exposure-dose and response well characterized • Contributions of mixtures, other environmental stressors, genetic variability in response and low-dose-response enabled	• Meta-analyses[1] of multiple epidemiological and clinical studies using molecular patterns associated with lung cancer in smokers, absent in nonsmokers • Marginal characterization of exposures and exposure-dose; for PAHs, human exposure was characterized by self-reported numbers of cigarettes smoked • Experimental measurement of dose-dependent, chemically induced alterations in transcriptomics in humans, using specific and sensitive assays • Variability of exposure-dose and response less well characterized • Evaluation of multiple BaP studies attempted in rodents, but study quality was inadequate
Strengths:	• Augment characterization of hazard and exposure-dose-response using molecular patterns • Better characterize associated or causal mechanisms of health effects from chemical exposures • Better describe population variability • Enable characterization of less well-studied chemicals with similar mechanisms • Data mining methods[2] to survey the literature for BaP/PAHs are significantly faster and less expensive than other approaches; evaluate most existing data	
Weaknesses:	• Currently, traditional data needed to anchor molecular estimates of risk • Currently, molecular epidemiology and clinical studies are neither faster nor less expensive than traditional approaches but are improvements over traditional data alone • Nonhuman data with nonconcordant tissue responses are challenging to extrapolate to humans • Much published molecular biology data is inadequate for risk assessment due to limitations in use of best practices, analyses, and reporting • Many sources of variability can lead to false associations	

[1]**Meta-analysis** methods that combine data or results from multiple independent studies that seek to test similar hypotheses (Ramasamy et al. 2008).
[2]**Data mining** attempts to discover useful patterns or relationships in large amounts of data using advanced statistical methods, such as cluster analysis, artificial intelligence, or neural network techniques.

3.1.1 Benzene-induced Leukemia

Benzene is among the 20 most widely used chemicals in the United States and one of the most common environmental contaminants. A component of crude oil and gasoline, benzene also is used as an intermediate in the manufacture of resins, dyes, chemical solvents, waxes, paints, glues, plastics, and synthetic rubber. The major sources of benzene exposure are anthropogenic and

include fixed industrial sources, fuel evaporation from gasoline filling stations, and automobile exhaust. Benzene has been measured in outdoor air at various locations in the United States at concentrations ranging from 0.02 ppb (0.06 μg/m^3) in a rural area to 112 ppb (356 μg/m^3) in an urban area (IARC 2012). Personal monitoring of benzene exposure in Detroit, Michigan reported a mean of 1.72 ppb (5.5 μg/m^3) (George et al. 2011). The maximum contaminant level in drinking water is 5.0 μg/L or 5 ppb (EPA 2013b). The Occupational Safety and Health Administration permissible exposure limit for benzene workers in the United States is 1 ppm (OSHA 2014).

Benzene is a known human hematotoxicant and carcinogen (ATSDR 2007; EPA 2000; IARC 2012; NIOSH 1992). Epidemiological studies have associated benzene exposure with an increased risk of acute myeloid leukemia (AML), myelodysplastic syndrome, hematotoxicity (toxicity to the blood), and other blood disorders (EPA 2000; Goldstein 1988; IARC 2012; Schnatter et al. 2012). AML is characterized by uncontrolled proliferation of clonal neoplastic cells and accumulation in the bone marrow, with an impaired differentiation program. AML accounts for about 30 percent of all adult leukemias and is the most common cause of leukemia death (Howlader et al. 2013). Studies indicate that benzene also might cause lymphoma and childhood leukemia (Smith, M. T. et al. 2011). The extensive molecular epidemiological and clinical data sets for benzene-induced hematotoxicity and leukemia are ideal for exploring how new data types might be used to inform risk assessments. The work described in this section focuses on studies in which traditional and molecular data were collected simultaneously using a variety of methods, including genome-wide analyses of susceptibility genes (using genomic methods), protein expression (proteomics), and epigenetic modification (epigenomics). The studies also were conducted over a range of environmental exposure levels (<0.1 ppm to ≤10 ppm). A systems biology analysis of benzene-induced hematotoxicity and leukemia is summarized in McHale et al. (2012) and Smith et al. (2011). The information presented in these reports was developed primarily by Martyn Smith and colleagues at the University of California, Berkeley.

3.1.1.1 Systems Biology of Benzene-induced Disease

Benzene is among the most well-studied environmental chemicals, yet our understanding of the molecular mechanisms underlying hematopoietic cancer is somewhat recent (see Box 4 for a brief description). In 2009, McHale et al. identified exposure-dependent alterations in the genes and pathways of peripheral blood mononuclear cells (using transcriptomics) and hematotoxicity associated with benzene exposure (>10 ppm) in occupationally exposed Chinese workers (McHale et al. 2009). McHale et al. (2010) extended these findings to lower exposure levels of <1 ppm to ≤10 ppm.[10] R. Thomas et al. subsequently demonstrated changes in gene expression in Chinese workers exposed to levels <0.1 ppm, that is, below current U.S. urban levels (Thomas, R. et al. 2014). The exposure-response models used in these analyses were not selected a priori; instead,

[10]The McHale et al. (2010) study included 250 benzene-exposed workers and 140 unexposed age- and sex-matched controls who worked in 3 clothes-manufacturing factories in the same region of China. Transcriptomic profiles for exposed and unexposed individuals and among four exposure groups were compared. Exposure groups were based on occupational surveys and individual urinary benzene biomarkers.

their selection was driven by the best fit of the data. Results are consistent with supralinear exposure-responses, which also have been reported in some traditional epidemiology studies (Lan et al. 2004).

Based on these and other studies, the systems biology of benzene-induced early effects has been summarized by McHale et al. (2012) and others (Smith, M. T. et al. 2011; Zhang, L. et al. 2010a). Benzene-induced hematotoxicity and leukemia are thought to be initiated when metabolites of benzene interact with genes or pathways in hematopoietic stem cells that are critical to hematopoiesis. Interactions among various cell types within the bone marrow and among various tissues also play a role in leukemia (e.g., immunosurveillance). Mechanisms of benzene-induced hematotoxicity and leukemia (shown in Figures 5 and 6, below) center on exposure-dependent pathway alterations comprising 147 significant genes altered in peripheral blood mononuclear cells from humans exposed to benzene (cross validated on two microarray test platforms [Illumina and Affymetrix] and ribonucleic acid [RNA] sequencing) (see below). The benzene-related gene expression profiles change with dose, with some genes (and related biological processes) expressed at all levels and others expressed only at higher concentrations. Of the 147 genes, the expression of 16 was significantly altered at all exposure levels. These 16 signature genes are involved in immune response, inflammatory response, cell adhesion, cell matrix adhesion, and blood coagulation, and are most strongly associated with AML pathways (McHale et al. 2010). This set of 16 genes can be used collectively as a biomarker[11] (or "gene signature") for chemical exposure to benzene-associated hematotoxicity. Given the strong evidence linking hematotoxicity in benzene-exposed populations to leukemia, this gene signature also is anticipated

Box 4. Molecular Mechanism of Acute Myeloid Leukemia (AML)

The probable mechanism by which benzene induces leukemia involves the "targeting of critical genes and pathways" (McHale et al. 2012). Benzene can induce abnormalities in the genes, chromosomes, or epigenetic mechanisms of hematopoietic stem cells (HSCs). Benzene also can disrupt the normal cell cycle, leading to apoptosis, increased cell proliferation, and altered differentiation of the HSCs. Benzene causes these effects and ultimately leukemia by inducing oxidative stress, dysregulating proteins that control normal functioning of HSCs, and reducing the body s ability to detect and destroy cancer cells (McHale et al. 2012).

Two events that are important for leukemic transformation have been identified. The first event is uncontrolled cell growth, which is mediated by upregulation of cell survival genes. The second is alteration of transcription factors that control HSC differentiation. That is, the genes that encode transcription factor proteins can be mutated or can target the expression of certain genes in a way that interferes with the appropriate differentiation of HSCs.

For AML specifically, two major types of genetic events have been described that are crucial for leukemic transformation. A proposed necessary first event is disordered cell growth and upregulation of cell survival genes. The most common of these activating events was observed in the receptor tyrosine kinase (RTK) *Flt3*, in the genes *N Ras*, *K Ras*, and *Kit*, and sporadically in other RTKs. Alterations in myeloid transcription factors governing hematopoietic differentiation provide the second necessary event for leukemogenesis. Transcription factor fusion proteins such as AML ETO, PML RAR alpha, or PLZF RAR alpha block myeloid cell differentiation by repressing target genes. In other cases, the genes encoding the transcription factors themselves are mutated. (Kanehisa Laboratories 2014a; Wang, I. et al. 2012a).

[11]**Biomarkers** are characteristics that are measured and evaluated objectively as indicators of normal biological processes, pathogenic processes, toxicological response to an environmental exposure, or pharmacological responses to an intervention. Adapted from Institute of Medicine (2010).

to be predictive of future leukemias in benzene exposed populations, and potentially for exposure to leukemogens in general. In a subsequent study, Thomas R. et al. (2013a) also evaluated benzene-related molecular changes using a different technology, RNA sequencing, and observed results generally consistent with the microarray results regarding benzene-induced changes.[12] The work of R. Thomas et al. (2014) and

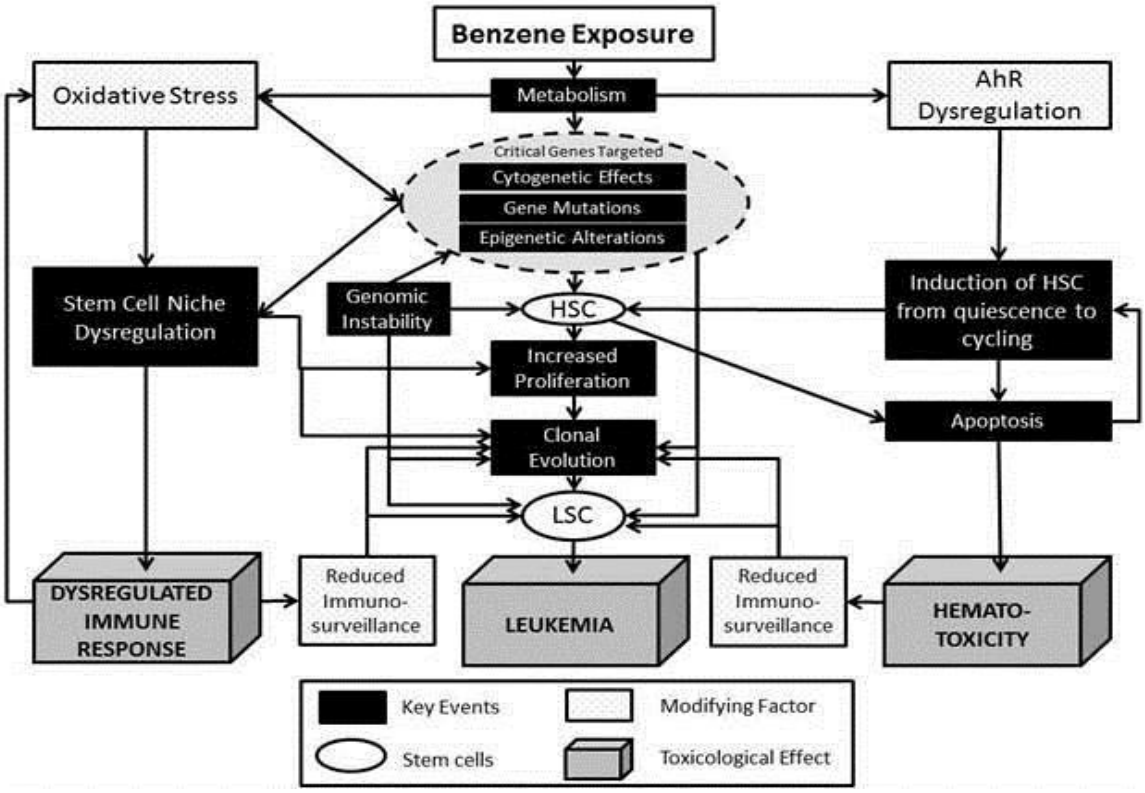


Figure 5. Multiple Modes of Action (MOAs) (also called Adverse Outcome Pathway (AOP) Network) for Benzene-induced Leukemogenesis.

The legend depicts potential key events, modifying factors, and toxicological effects. Stem cells can be either HSCs (hematopoietic stem cells) or LSCs (leukemic stem cells) (Smith, M. T. et al. 2011). The figure also highlights mechanistic commonalities with other chemical leukemogens and idiopathic leukemia (i.e., unknown or spontaneous origin). Reproduced with permission from *Elsevier*.

---

[12]"The Pearson correlation between the two technical replicates for the RNA-seq experiments was 0.98 and the correlation between RNA-seq and microarray signals for the 20 subjects was around 0.6. Sixty percent of the transcripts with detected reads from the RNA-seq experiments did not have corresponding probes on the microarrays. Fifty-three percent of the transcripts detected by RNA-seq and 99% of those with probes on the microarray were protein-coding. There was a significant overlap ($P < 0.05$) in transcripts declared differentially expressed due to benzene exposure using the two technologies. About 20% of the transcripts declared differentially expressed using the RNA-seq data were noncoding transcripts. Six transcripts were determined (false-discovery rate <0.05) to be alternatively spliced as a result of benzene exposure. Overall, this pilot study shows that RNA-seq can complement the information obtained by microarray in the analysis of changes in transcript expression from chemical exposures" (2013a).

McHale et al. (2010) exemplifies how such biomarkers could be used, particularly in augmenting traditional epidemiology studies and enabling new types of molecular epidemiology studies at lower concentrations.

Exposure to benzene also induces a distinct lymphoma disease signature (McHale et al. 2010; McHale et al. 2012; Smith, M. T. et al. 2011). The traditional epidemiological data on lymphoma are inconclusive. Characterization of a benzene-induced molecular mechanism for lymphoma adds considerably to the evidence for benzene-induced lymphoma. This characterization is a good example of using molecular mechanistic data to support the MOA and to strengthen the evidence determinations (IARC 2012).

One important caveat regarding individual epidemiology studies is that they provide evidence of association *not* causality. Establishing causality requires meta-analyses of multiple, well-conducted epidemiology studies, experimental data from clinical or animal studies, or mechanistic understanding (EPA 2005, 2009a; U.S. DHHS 2014). For the benzene prototype, data from multiple epidemiological studies and mechanistic information from multiple sources (Kanehisa Laboratories 2014b) were used. The causal relationships between specific gene/pathway alterations and leukemia are best supported by clinical studies using chemotherapeutic agents that alter expression of specific genes in the critical pathways with results that demonstrate either the blocking or amelioration of idiopathic disease (i.e., unknown or spontaneous origin) outcomes (Hatzimichael and Crook 2013).

3.1.1.2 Idiopathic and Other Chemical Leukemogen-induced Disease

Molecular mechanisms for benzene-induced leukemia appear similar to idiopathic AML, as well as AML induced by other environmental agents (e.g., alkylating agents, topoisomerase II inhibitors) (IARC 2012; McHale et al. 2012; Pedersen-Bjergaard et al. 2008). Figure 6[13] shows a network of genes and pathways thought to be causally related to both idiopathic and chemically induced leukemia (NIH BioSystems, Kyoto Encyclopedia of Genes and Genomes (KEGG); Kanehisa Laboratories 2014a). Note that this diagram illustrates only a subset of the complete set of processes involved in AML (see NIH BioSystems; Kanehisa Laboratories 2014b; McHale et al. 2010; 2011; Thomas, R. et al. 2014). The circles in the figure indicate some of the specific genes and pathways affected by leukemogenic agents and environmental modifiers (IARC 2012; Kanehisa Laboratories 2014a; McHale et al. 2010; Pedersen-Bjergaard et al. 2008). Additional evidence for the causal role of these genes and pathways in AML is provided by the study of human genetic variants associated with altered risks and chemotherapeutics that reverse adverse alterations in some of these same genes and pathways (discussed below). Although mechanistically similar, different agents can display specific characteristics such as origins in cells at different stages of

[13]The basic AML network figure used in Figures 6 and 7 is from the Kyoto Encyclopedia of Genes and Genomes (KEGG; Kanehisa Laboratories 2014a); also reported in National Institutes of Health BioSystems database. The added circles are the work of the report authors.

hematopoiesis, distinct cytogenetic subtypes, and different latencies (Irons et al. 2013; McHale et al. 2012).

Figure 6 highlights how a network of related events can be modified at different points but still lead to a common disease outcome. These mechanistic commonalities and differences among idiopathic and chemically induced health effects can be used to characterize chemicals with limited data. In other words, data-limited chemicals would be of elevated concern if they alter pathways similar to what is observed in idiopathic disease or with well-studied leukemogens. For example, R. Thomas et al. (2012a) used existing information on gene and protein targets of 29 known leukemia-causing chemicals and 11 carcinogens that are not known to cause leukemia. The authors were able to develop a classification scheme that could distinguish a random leukemia-causing/nonleukemia-causing carcinogen pair with 76 percent probability. Later in this section, additional support for the similarity of mechanisms for chemical-related and idiopathic diseases is provided (see the ozone and B[a]P prototypes). These examples highlight how mechanistic information improves our ability to understand and assess cumulative risks.

3.1.1.3 Cumulative Risks from Environmental Factors

New approaches can help characterize cumulative contributions to potential risks for disease from various environmental factors, including exposure to chemicals. Evidence suggests that, in addition to environmental exposures, genetic variations and lifestyle factors such as smoking, obesity, diet, and alcohol use are risk factors for leukemia (Belson et al. 2007; Ilhan et al. 2006; Pedersen-Bjergaard et al. 2008; Smith, M. T. et al. 2011). Figure 6 shows how multiple environmental factors can alter various molecular events in a way that is likely to alter risks for a specific disease. The figure also illustrates how chemicals might be included or excluded based on a common mechanism and potential contribution to cumulative risks. Evaluating exposures to the developing organism as a potential risk factor for disease later in life also is important, especially because of the potential of benzene and other environmental agents to alter epigenetics in the developing organism (which is highly sensitive to epigenomic changes), as well as the association between environmental exposures to benzene and childhood leukemias (Boekelheide et al. 2012).

Individuals exposed to known environmental and lifestyle risk factors account for only approximately 20 percent of the acute leukemia incidences, indicating that host genetic susceptibility might be a key factor in onset of disease (Smith, M. T. et al. 2011). These new approaches could dramatically improve our ability to characterize the potential disease susceptibility of subpopulations by distinguishing the extent to which chemicals, nonchemical stressors, and intrinsic genetic variations[14] contribute to alterations in the same genes and biological pathways. Genetic variation is discussed more specifically below, and an example of altered subpopulation risks based on genetic variations is provided.

---

[14]Human genetic variation is evaluated by identifying genetic differences among subpopulations. Multiple variants of any given gene can occur in the population. These differing DNA codings determine distinct traits or polymorphisms that can influence risks.

3.1.1.4 Genetic Variation and Susceptibility in the Human Population

New approaches are improving our ability to characterize genetic variation and susceptibility to both idiopathic and chemically induced disease. For example, several genetic variations appear to increase risks for developing AML, while at least one decreases risks (Garte et al. 2008; North et al. 2011; Shen et al. 2011; Smith, M. T. et al. 2011; Zhuo et al. 2012). Sillé et al. (2012) reported 12 independent risk loci with the potential to alter gene expression related to AML. Independent risk loci are specific regions within the genome, which can be a single base, as in this case, or an entire gene. A significant number of variants (i.e., single nucleotide polymorphisms [SNPs] related to a tumor suppressor gene, signaling pathways, or residing in putative regulatory elements)[15] have been linked to different types of multiple hematological cancers. Figure 7 highlights genes that vary in the human population and are associated with altered leukemia risk. Chemotherapeutic agents that change these "implicated" genes to a more normal state also decrease the incidence of leukemia, providing supporting evidence that these genes and pathways are involved in the disease process (Kanehisa Laboratories 2014a). Figure 8 presents the results of a meta-analysis of epidemiological data on the differential risks for acute leukemia associated with one human variant. The individual epidemiological study results and the pooled results are shown. In this case, a SNP leads to a substitution of isoleucine with valine at codon 462 in exon7 (Ile462Val or CYP1A1*2C polymorphism, rs1048943). This exon7 polymorphism results in three genotypes: a predominant homozygous Ile/Ile, the heterozygote Ile/Val, and a rare homozygous Val/Val. The overall risk was 42 percent greater (95% CI = 1.11–1.98) for the Val/Val plus Val/Ile genotypes versus the Ile/Ile CYP1A1 genotype (Zhuo et al. 2012). An alternative hypothesis is that this SNP is not causative but rather is linked to a causative SNP not identified in the epidemiological study.

Characterizing the potential susceptibility of subpopulations to disease incidence due to individual genes, combinations of genes, and gene variants can be very challenging, as many genes can interact to alter susceptibility. Although subpopulations can be categorized according to variant profile and susceptibility, individual risk is likely influenced by a variety of factors including individual genomics and epigenomics. Section 4 details NexGen approaches that might substantially improve our ability to characterize human susceptibility and estimate the contribution of various risk factors (e.g., lifestyle, genetic variability, exposure to chemicals) to overall risk.

3.1.1.5 Benzene *In Vitro* Evaluation of Toxicogenomic Signatures

How well *in vitro* assays predict *in vivo* outcomes or otherwise inform our understanding of chemical risks is a topic of great interest and the subject of much active research. Godderis et al. (2012) conducted an *in vitro* study of benzene in a human lymphoblastoid cell line (TK6) to detect gene signatures and biological pathway perturbations. The global gene expression resulting from exposure to 15 genotoxic carcinogens, including benzene and its metabolites, was evaluated. The goal was to determine if well-characterized chemicals could be used to characterize data-limited

[15] **Putative regulatory elements** are areas of the gene that do not code for proteins but rather regulate DNA expression via transcription into proteins.

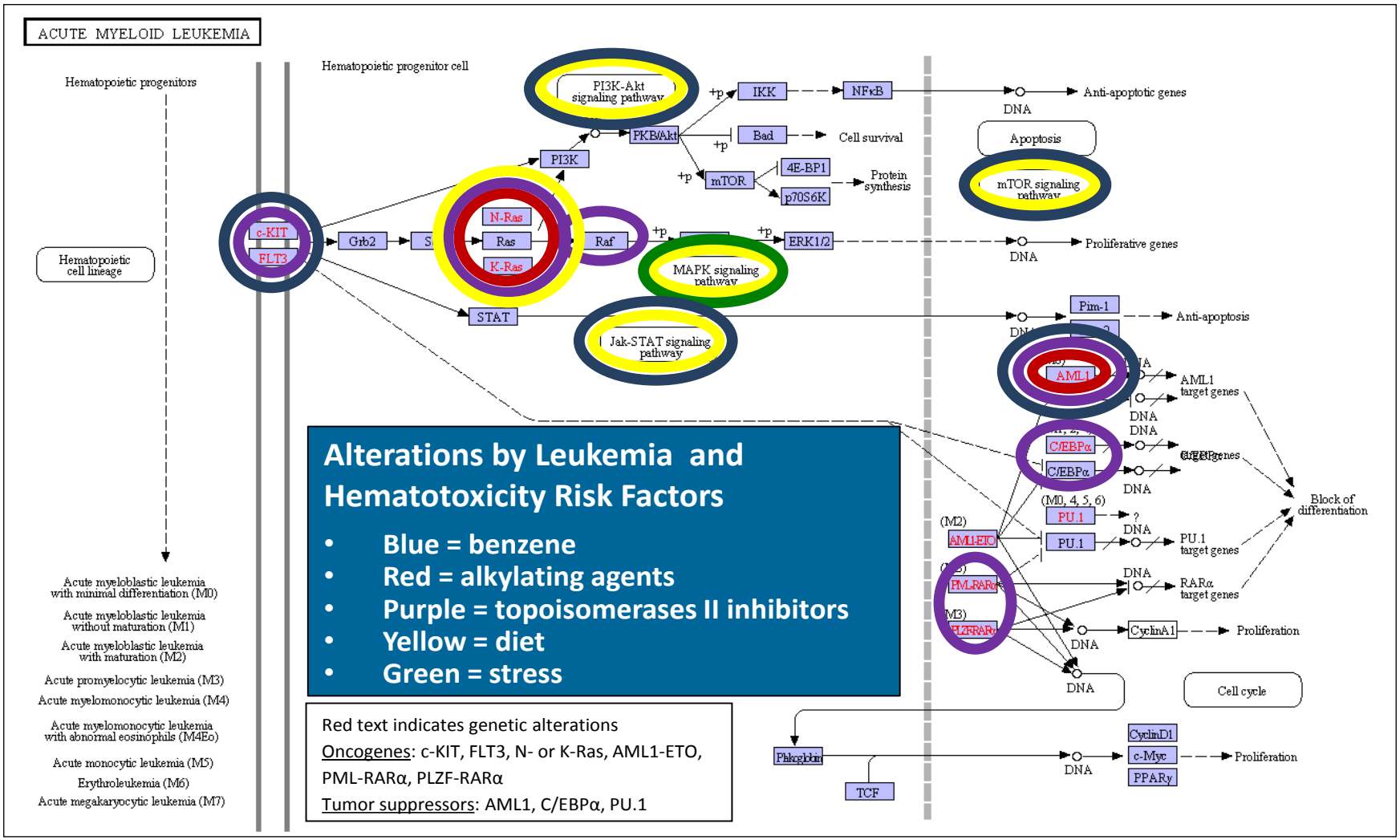


Figure 6. The Kyoto Encyclopedia of Genes and Genomes (KEGG) Diagram with Chemical and Lifestyle-induced Alterations Shown.
Some of the currently understood molecular pathways involved in acute myeloid leukemia (AML). Altered oncogenes and tumor suppressor genes are noted in red type (Kanehisa Laboratories 2014a). The circles (added by authors) note specific genes and pathways that are modified by benzene, other chemical leukemogens, and other risk factors. The solid vertical lines indicate the cell membrane and the dashed vertical line indicates the nuclear membrane. This diagram is intended to be illustrative rather than comprehensive, but it shows how single or combinations of environmental factors could modify risks for leukemia, and how such knowledge could be used to evaluate joint effects of environmental factors (IARC 2012; McHale et al. 2012; Pedersen-Bjergaard et al. 2008; Smith, M. T. et al. 2011).

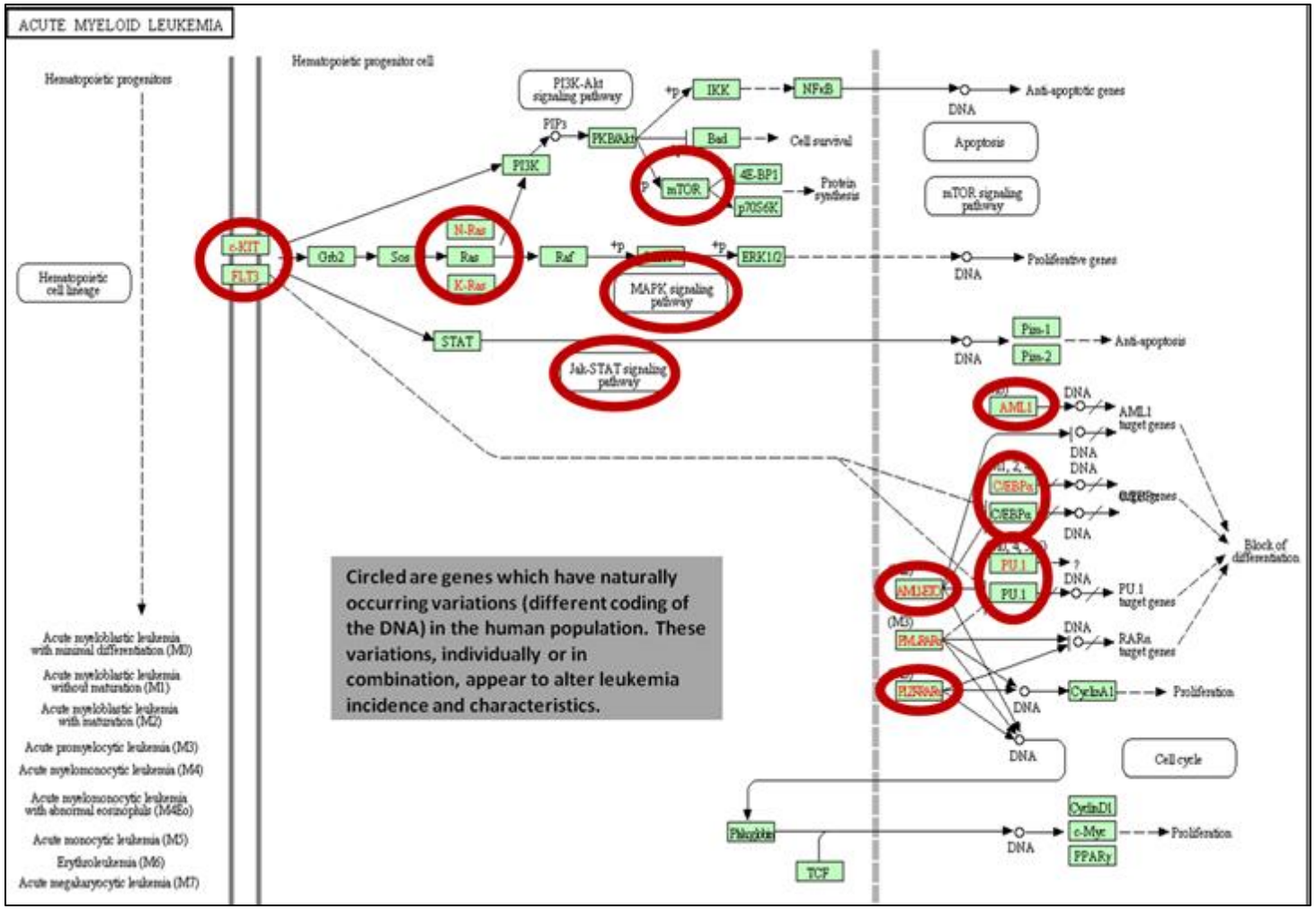


Figure 7. The Same Kyoto Encyclopedia of Genes and Genomes (KEGG) Diagram for Acute Myeloid Leukemia (AML) as Shown in Figure 6, with Human Gene Variants Circled.

In this version, the circles indicate the locations of naturally occurring human genomic variants that increase the risk of AML (Hatzimichael and Crook 2013; Sille et al. 2012). Characterizing genomic variant subpopulations and associated risks can help better describe human variability and susceptibility for specific diseases. Circles added by authors.

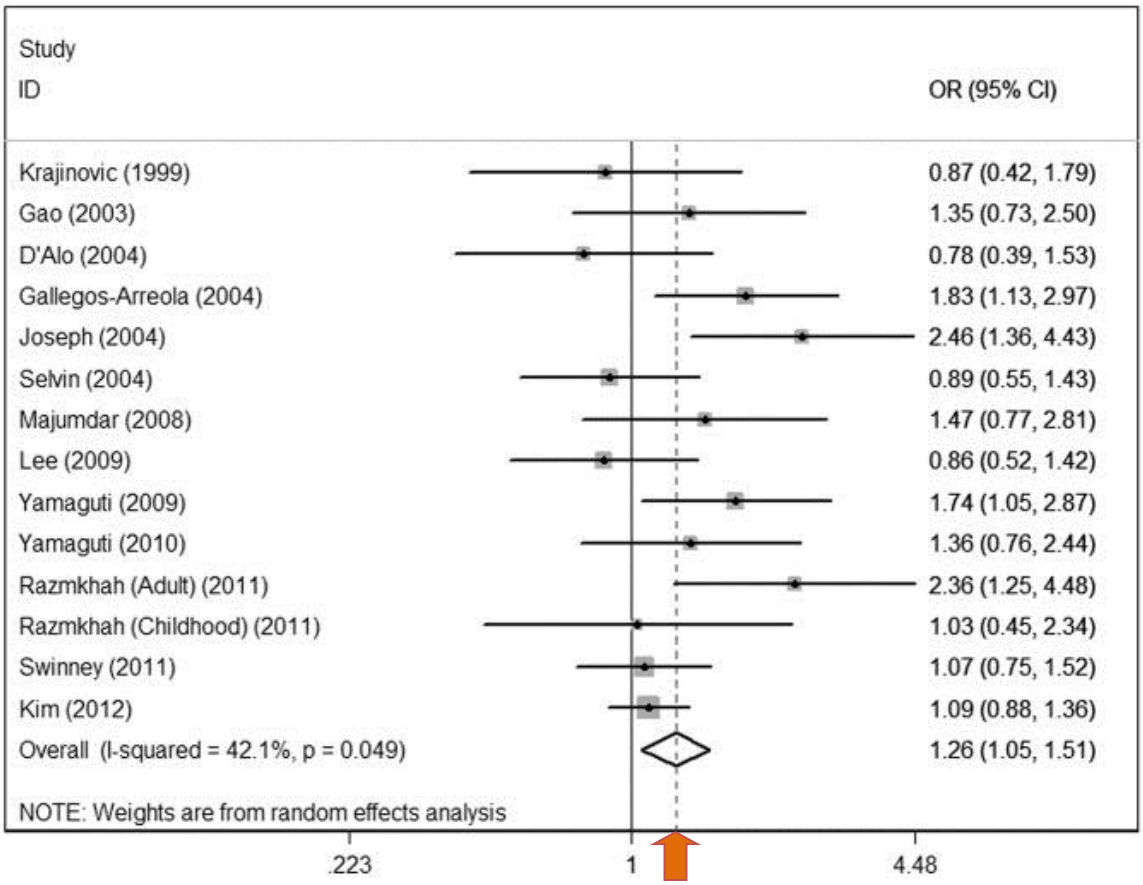


Figure 8. Meta-analysis for the Association of Acute Leukemia Risk with CYP1A1 Ile462Val Polymorphism. (OR = odds ratio). The overall risk was 42 percent greater (95% CI = 1.11–1.98) for Val/Val + Val/Ile versus Ile/Ile (Zhuo et al. 2012). Reproduced with permission from *PLoS One*.

chemicals by comparing gene signatures. Although results on pathways altered by exposure to benzene and its metabolites were in general agreement with those shown in previous *in vivo* studies, this was not generally true for most of the chemicals evaluated. The authors pointed out that several factors could complicate comparison of *in vivo* and *in vitro* data. For example, metabolism is limited in the *in vitro* systems, and the addition of metabolic enzymes (e.g., S9) had confounding effects. Responses also can differ depending on cell type, confounding comparisons between outcomes for various *in vitro* cell lines and *in vivo* results. The authors supported the use of toxicogenomic signatures for evaluating data-limited chemicals, but for the carcinogens in this study, they were unable to determine discriminatory mechanisms based on *in vitro* data alone. This suggests that developing putative mechanisms of action based on meta-analyses of human disease combined with mapping *in vitro* data against this information might prove more successful than attempting to understand mechanisms of action based solely on *in vitro* data. Considerable additional work will be necessary to develop the most efficient and appropriate mix of *in vitro* and *in vivo* data and methods for application of this approach to large numbers of chemicals and diseases.

3.1.1.6 Risk Assessment Implications Based on the Benzene Prototype: Use of New Data

The benzene prototype demonstrated the feasibility of using molecular biology data, particularly mechanistic signatures, in hazard identification and exposure-dose-response assessment.

Hazard Identification

Genes and pathways altered by benzene exposures are strongly associated with a network of pathways thought to be causative for known (hematotoxicity and AML) and likely (lymphoma) outcomes. The benzene results have been reproduced in multiple experiments using two different microarray assay platforms and an alternative technology, RNA sequencing. Evidence for a causal relationship between alterations in specific gene pathways, hematotoxicity, and leukemia risks is provided by observed similarities in pathway disruptions caused by other chemical leukemogens or observed in leukemia of unknown origins. A decreased incidence or severity of the disease by certain leukemia chemotherapeutic agents that reverse these adverse pathway changes provides further support. This ability to alter pathways biochemically and change the risk or attributes of the disease provides strong experimental evidence of the causal nature of gene/pathway alteration in leukemia. The molecular epidemiology and molecular clinical study data provide further evidence that gene signatures can be used to predict specific diseases with some confidence. Thus, well-defined pathway and network disruptions, strongly associated with a specific disease, could be informative in risk assessment for hazard identification and for low-dose response characterization. A well-characterized AOP also might provide context to interpret high-throughput data for many chemicals that do not have traditional data, assuming that chemicals that induce comparable effects on sufficiently understood pathway mechanisms likely would increase risks for the same disease outcome.

Exposure-Dose-Response Assessment

Increasing the benzene dose resulted in significant dose-dependent alterations in gene transcription. Some genes that were associated with cytotoxicity and cell death were transcribed only at higher exposures. A specific 16-gene signature, observed at all environmental exposure concentrations measured (<0.1 to >10 ppm), was identified and was associated generally with altered immune function, hematotoxicity, and leukemia. This signature can serve as an indicator (or biomarker) of both exposure and effect. The exposure-response models used to describe the data were not specified in advance, rather they represented the best fit from among multiple models. Hence, the model was "agnostic" on the issues of threshold/no threshold and nonlinear/linear. For the 16-gene signature, the statistical best fit for the exposure-dose-response relationship at environmental concentrations was linear. No threshold was observed (Thomas, R. et al. 2014).

The above discussion highlights several ways these data and approaches can improve exposure-dose-response assessment by:

- providing tools to study the complex interactions among pathways that help organisms adapt to insults or increased risks;
- identifying dose-dependent molecular biomarkers that can be used to characterize exposure-dose-response relationships in the range of environmental exposures, replacing

estimates based on extrapolating from higher dose empirical data (assuming the more traditional studies have demonstrated the power to detect potential responses); and

- using models that best fit the relevant empirical data to reduce uncertainty concerning model choice for extrapolations.

Molecular-based signatures (or biomarkers) are anticipated to become more common in the future. Such biomarkers can be used to develop data on more environmentally relevant exposure levels than currently available from traditional epidemiological studies and to reduce measurement error. When calibrated to known outcomes, molecular-based signatures could be used to measure the exposure-response relationship directly in the human population, similar to how simpler biomarkers are currently used to quantify lead exposures and effects (Mendrick 2011).

Cumulative Risk Assessments

Interpreting chemically induced events within the context of an already characterized disease mechanism illustrates how chemicals can affect a network of related pathways at multiple points. Chemicals that increase risks for the same disease might not have the same molecular target(s). An illustration of how known chemical leukemogens and risk factors for leukemia alter different pathways in a network of events associated with hematotoxicity and leukemia was presented in Figure 6 (in Section 3.1.1.4). Integrating chemical effects at this network level demonstrates how one might account for the contribution that various chemicals or other environmental stressors and factors might have to an overall cumulative risk. The benzene prototype is a good example of how sufficient mechanistic knowledge can facilitate cumulative risk assessment. It also demonstrates how caution is warranted for the predictive capability of overly simplified descriptions of an AOP network that do not support accurate estimates of the cumulative risks from chemical exposure and other critical disease factors.

Intraspecies Variability and Population Response Distributions

The benzene prototype demonstrates improvements in characterizing subpopulation responses due to genetic variability. Specific genes were identified for which variants in the human population are associated with altered incidence of, and prognosis for, leukemia. An example is also provided of how variants of a single gene are associated with altered relative risks for the variant subpopulations, although a comprehensive analysis of the mechanistic linkage for this association is still needed. As additional research and data evolve from personalized medicine, our understanding of human variability in disease response to chemical exposure could be significantly improved. Data-driven characterization of human variability and population response distributions would improve both cancer and noncancer risk assessments, and lead to a more harmonized approach.[16]

In summary, the benzene prototype exemplifies how toxicogenomic data from environmental exposure in humans can be used to improve our mechanistic understanding of the onset of disease,

[16]Current methods to estimate risks and account for human variability differ for cancer versus noncancer responses because of a lack of empirical data characterizing targets and the mechanisms leading to disease.

the ability to better estimate cumulative risks and identify susceptible subpopulations, and characterization and estimates of the low dose-response relationship; all of these are historically challenging issues in risk assessment.

3.1.2 Ozone-induced Lung Inflammation and Injury

Hundreds of controlled human exposure studies have described biological changes in volunteers exposed acutely (usually for 2–6 hours) to ozone concentrations ranging from 0.06 to 0.4 ppm and have documented the relationship between ozone exposure and inflammation (EPA 2013c).[17] These studies demonstrate that exposure to ozone causes decrements in lung function, increases in markers of pulmonary inflammation and lung injury, and alters host defenses against inhaled pathogens. The data on ozone represent the single largest human clinical database of any pollutant EPA has studied. Inflammatory responses resulting from acute exposures are of public health concern. As a consequence, and because the mechanisms are well understood, this *in vivo* database provides an ideal opportunity to demonstrate proof of concept for using molecular biology and *in vitro* data to develop faster, more efficient approaches to assessing human health risks, following exposure to a toxicant (ozone) that induces oxidative stress (lung inflammation) and causes an inflammatory response.

Chronic inflammation is implicated in the etiology of several diseases, including atherosclerosis, heart disease, obesity, diabetes, arthritis, cancer, and lung diseases (asthma, emphysema, pulmonary fibrosis). Both common and disease-specific inflammatory molecular patterns have been reported to underlie these diseases (Wang, I. et al. 2012a). Why a particular disease is expressed in an individual or a subpopulation as the result of chronically induced inflammation likely depends on several factors, including the injury site, co-activation of other networks, genetic variation, or other environmental exposures. Such complicating factors highlight several challenges in predicting disease risks based on patterns of molecular changes. Nonetheless, observing an inflammatory signature for a chemical that has not been well studied likely would raise concerns for potential inflammatory disease risks. The specific inflammatory disease in question likely would be difficult to predict, however, if the systems biology context were limited. Any given network also might be involved in multiple disease outcomes. Conversely, a specific disease outcome could involve multiple interactive pathways and networks. If, however, chemicals with an inflammatory molecular signature were inhaled, it would be reasonable to assume that these chemicals could cause lung inflammation and injury.

3.1.2.1 Systems Biology Approach for Ozone-induced Lung Inflammation and Injury

The perturbation of a biological pathway initiates events that cause an adverse outcome associated with an environmental stressor. These perturbations must be evaluated for severity and distinguished from adaptive or associative pathway alterations. The proposed physiological and

[17]The current ozone standard calls for limitation of the fourth highest daily maximal 8-hour ozone concentration in a year to 0.075 ppm, based on a 3-year average.

cellular pathways by which ozone causes pathophysiological changes in the human respiratory tract are illustrated in Figure 9. The data and methods exemplified in this prototype focus on the pathways that lead to inflammation, which are shown in the open boxes in Figure 9 (see Box 5 for a description of inflammation). Alveolar macrophages and epithelial cells lining the respiratory tract are thought to be the primary lung cells responsible for inducing an inflammatory response. Lung epithelial cells are at least 100 times more abundant than alveolar macrophages and produce pro-inflammatory cytokines such as interleukin-8 (IL-8), which is a potential neutrophil chemoattractant. This project therefore focused on the response of epithelial cells to ozone. Pathways based on neurological responses to ozone exposure (e.g., lung function decrements) might be more difficult to characterize using *in vitro* approaches. An extensive review of the MOA (termed AOP network in this report) for ozone is found in the *Integrated Science Assessment for Ozone and Related Photochemical Oxidants* (EPA 2013c).

Box 5. Inflammation

Inflammation is the immune system's response to cell and organ damage by pathogens, chemicals, or physical insult. Initially, various inflammatory cells (e.g., neutrophils, lymphocytes) accumulate at the injury site. Cell debris resulting from the lung injury or pathogens is removed as tissues begin to repair. If the balance between inflammation and resolution of the events leading to the inflammation is dysregulated, or tissue insult continues, inflammation can lead to disease pathology (Medzhitov 2008; Wang, I. et al. 2012a).

Understanding adverse *in vivo* outcomes in terms of perturbations to normal biological pathways identified with a set of *in vitro* assays would enable the results of these assays to be used to build qualitative or quantitative models of chemical-biological activity relationships that could predict *in vivo* responses based on *in vitro* data. For *in vitro* pathway information to be used quantitatively in risk assessment, the relationship between perturbation of a pathway following *in vitro* exposure and downstream endpoints (i.e., pathophysiological changes at the tissue or organism level following *in vivo* exposure to animals or humans) must be established. Establishing such a quantitative relationship currently is not possible for most toxicants that EPA is responsible for regulating because of insufficient *in vivo* and *in vitro* data (Crump et al. 2010a). Because several human studies characterize inflammation at multiple ozone concentrations and times after exposure, this rich data set of human *in vivo* responses can be used to investigate associations with *in vitro* assay results within the context of an AOP.

3.1.2.2 Primary Molecular Events in the AOP Network for Ozone-induced Inflammation (Step 1 in Figure 9)

Many pollutants induce intracellular oxidative stress, which can affect signaling pathways and ultimately lead to activation of pro-inflammatory genes.[18] Until recently, whether ozone induced intracellular reactive oxygen species (ROS) was unknown. Figure 10 shows that ozone can induce a rapid dose- and time-dependent increase in cytosolic intracellular glutathione redox potential, a measure of ROS (Gibbs-Flournoy et al. 2013).

[18]Differences in pollutant specific outcomes, among pollutants which act via ROS, can arise from differences in delivered dose, site of ROS production, cell and tissues specific differences, and/or interaction with other pathways that contributed to health outcomes.

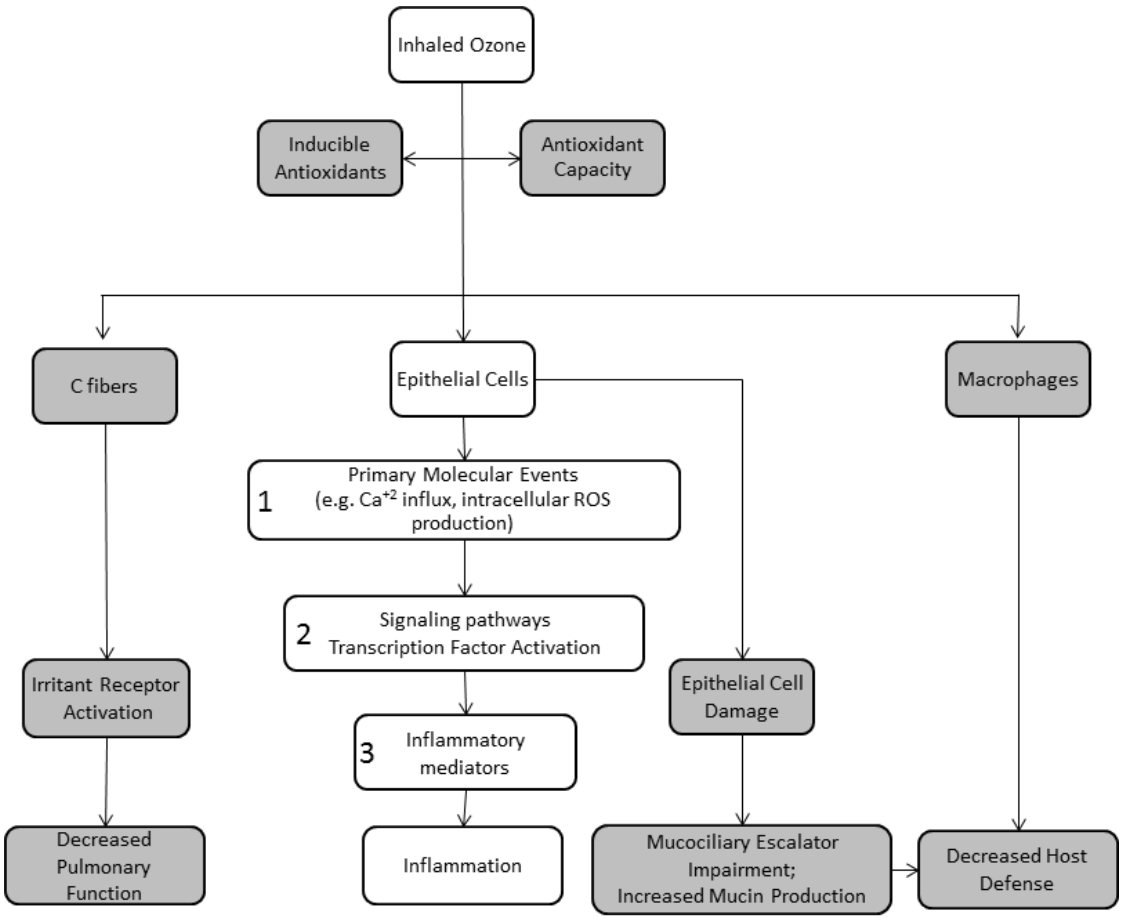


Figure 9. Proposed Key Events in Ozone's Modes of Action (MOA) *In Vivo*.
The data and methods exemplified in this prototype focus on the pathways that lead to inflammation (open boxes). Numbers in boxes tie boxes to descriptive text below.

3.1.2.3 Downstream Signaling Pathways Induced by Ozone (Step 2 in Figure 9)

Figure 11 shows potential signaling pathways by which ozone could activate pro-inflammatory genes. Previous studies have reported that the production of pro-inflammatory mediators by lung epithelial cells is mediated by the NF-κB signaling pathway, shown on the right side of Figure 11 (Jaspers et al. 1997; Wu et al. 2011). These experiments were conducted, however, using transformed or immortalized cell lines. A significant potential limitation of using *in vitro* approaches to predict *in vivo* responses is whether cell lines, which are the most common cells used for *in vitro* studies, respond to toxicants in the same way as primary cells removed directly from a host. Indeed, a recent study (McCullough et al. in press) shows that ozone-induced inflammation is not induced by NF-κB pathways in primary human airway epithelial cells (Figure 12). Rather it is induced by ERK and p38 pathways (Figure 13). (Pathways are shown on the right and middle portions of the AOP network diagram in Figure 11). These data underscore the importance of choosing appropriate cells for *in vitro* studies because, in this case, something as basic as how a cell regulates its ability to produce inflammatory mediators differs between primary cells and cell lines.

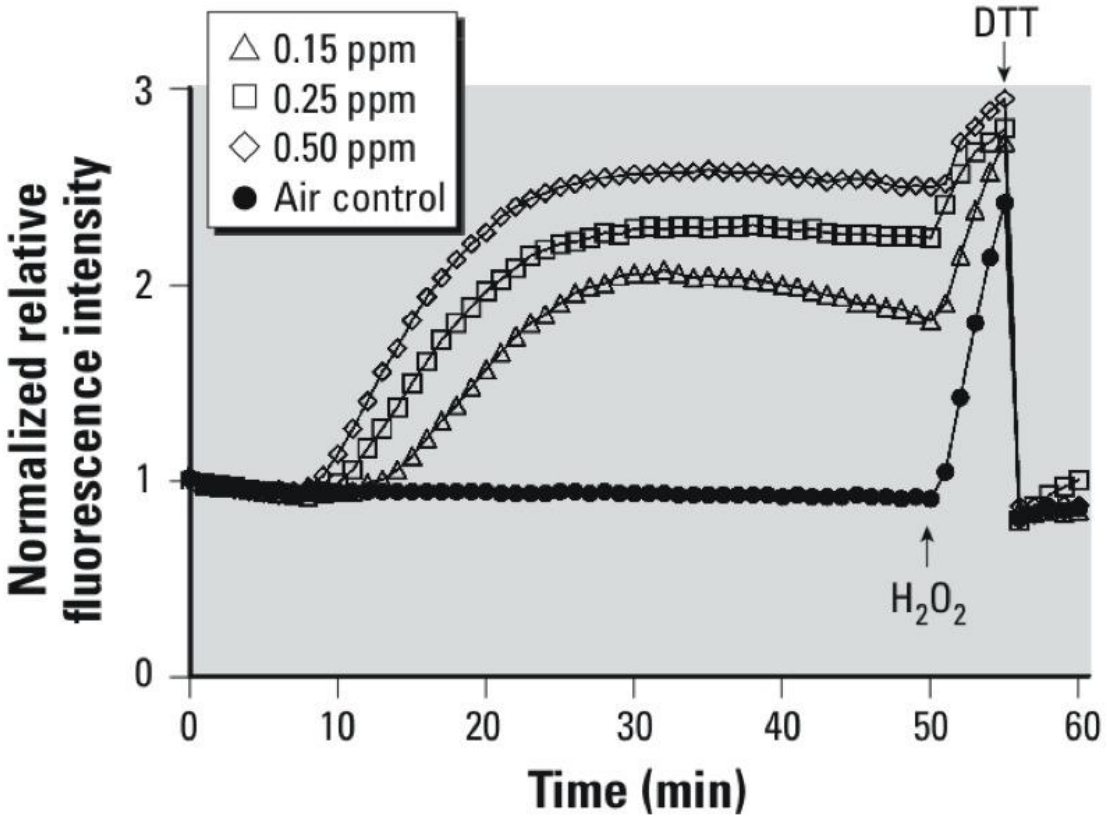


Figure 10. Exposure to Ozone Induces a Rapid Increase in Intracellular Reactive Oxygen Species (ROS). Addition of 0.1 mM H_2O_2 at the end of the ozone exposure produced a maximal response, which was fully reversible with the addition of 10 mM dithiothreitol, a strong reducing agent (Gibbs-Flournoy et al. 2013). Reproduced with permission from *Environmental Health Perspectives.*

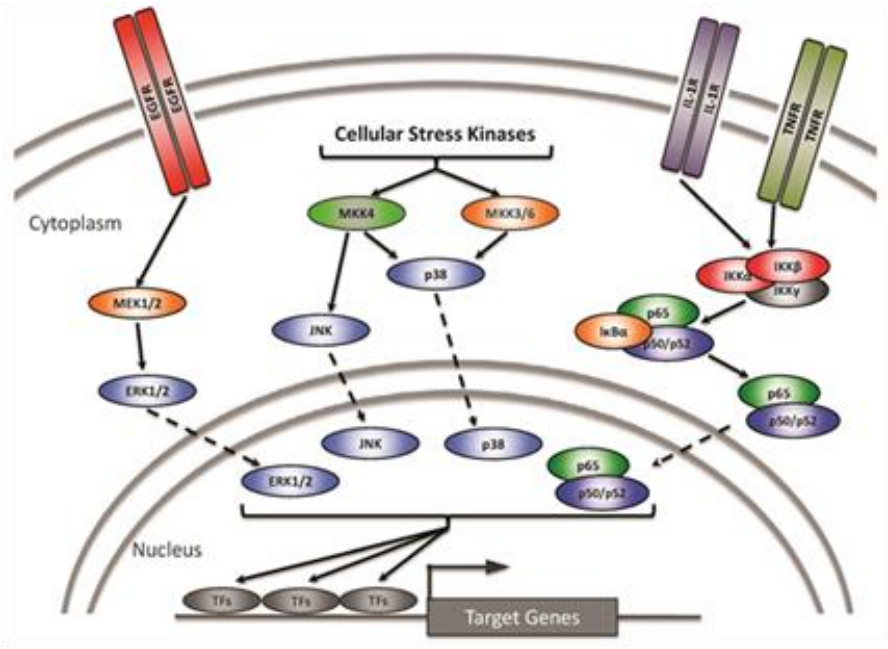


Figure 11. Potential Pathways by which Ozone Causes Production of Pro-inflammatory Mediators in Epithelial Cells.

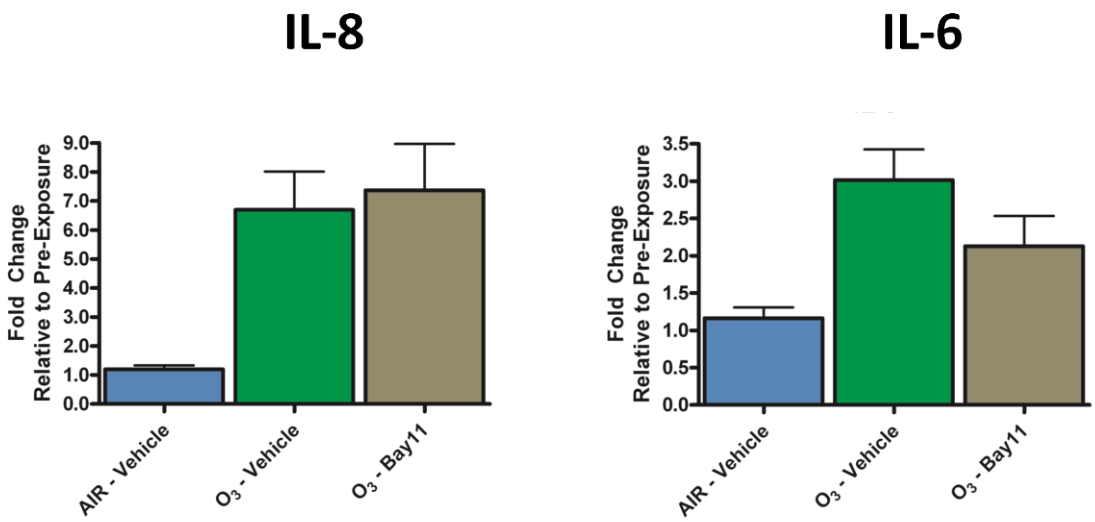


Figure 12. Ozone-induced Production of Pro-inflammatory Cytokines Interleukin 8 and Interleukin 6 Is Not Diminished When the NF-κB Pathway Is Inhibited by Bay11. (McCullough et al. in press).
(Adapted with permission from the *American Thoracic Society.* Copyright © 2014 *American Thoracic Society.* Official Journal of the American Thoracic Society.)

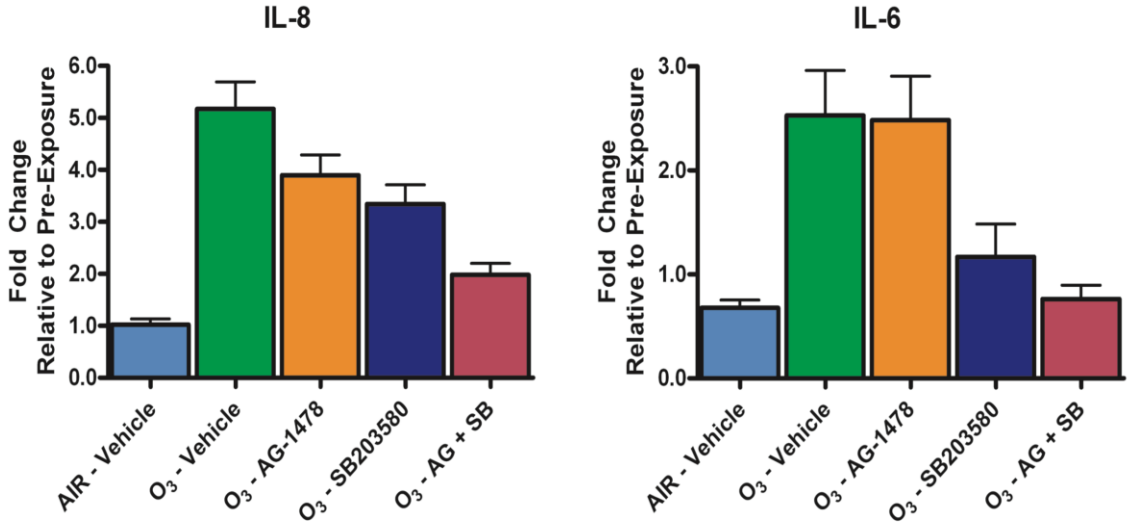


Figure 13. Ozone-induced Production of Pro-inflammatory Cytokines IL8 and IL6 Is Greatly Reduced When the ERK (Orange Bars) and p38 (Blue Bars) Are Inhibited (McCullough et al. in press).
Using small molecule kinase inhibitors, ERK and IL-8 are inhibited by 1μM AG-1478 (Orange Bars), p38, IL-8 and IL-6 are inhibited by 10μM SB203580 (Blue Bars), and ERK and p38, along with IL-8 and IL-6, are inhibited by a combination of 1μM AG-1478 and 10μM SB203580 (Red Bars). Also shown are clean air (Light Blue Bar) or 0.5ppm O_3 (Green Bar) both with dimethyl sulfoxide (DMSO) vehicle. (Adapted with permission from the *American Thoracic Society*. Copyright © 2014 *American Thoracic Society*. Official Journal of the American Thoracic Society.)

3.1.2.4 Characterization of Inflammatory Pathways Following *In Vivo* and *In Vitro* Exposure to Ozone (Step 3 in Figure 9)

Defining upstream events that control expression of inflammation is essential to characterizing the AOP by which ozone induces inflammation. Equally important is whether downstream events, such as the expression of pro-inflammatory genes following *in vitro* exposure to ozone, are consistent with the expression of those genes following *in vivo* exposure. To address the latter, a study was performed as outlined in Figure 14. Young, healthy volunteers were exposed to filtered air and a concentration of ozone (0.30 ppm) previously shown to induce a robust inflammatory response in the lung, including the production of pro-inflammatory cytokines such as IL-8 and IL-6. Bronchoscopy was used to obtain cells and lung fluid at 1 and 24 hours after each exposure. To ensure that pathophysiological effects observed in this study were comparable to those reported in earlier studies, downstream biomarkers of inflammation, such as the influx of neutrophils and production of pro-inflammatory cytokines (e.g., IL-8, IL-6) were measured (Devlin et al. 2012). Markers of cell injury (lactate dehydrogenase) and leakage of plasma components across the damaged epithelial cell barrier (albumin) into the lung airways were also measured. Bronchial airway epithelial cells were obtained by brush scraping, and microarray technology was used to define pathways affected by *in vivo* ozone exposure. In addition, quantitative proteomics was used to correlate changes in messenger ribonucleic acid (mRNA) measured by microarray with changes in their protein counterparts.

A subset of airway epithelial cells, collected from volunteers following exposure to filtered air, was cultured at an air-liquid interface. These cells were exposed to a tenfold range in ozone concentration and material collected for analysis at four time points after exposure. Similar to the *in vivo* studies, microarray and proteomics were used to identify and define pathways affected by ozone in these cells. Using cells from individuals exposed *in vivo* for *in vitro* exposures makes comparisons of *in vitro* and *in vivo* response from the same person possible. To identify an *in vitro* ozone concentration that is comparable to the *in vivo* concentration used, ozone with the heavy oxygen isotope (^{18}O) was used for both exposures. When ozone interacts with a target, the ^{18}O tag is deposited and can be measured by mass spectrometry. This ensures that cells are exposed to comparable *in vitro* and *in vivo* ozone doses. These experiments are described in Hatch et al. (in press).

Analysis of *in vitro* microarray data showed a concentration-dependent increase in the number of genes for which ozone altered the expression. At the lower concentrations, nearly all differential gene expression was upregulated, but at the highest concentration (1.0 ppm), many genes also were downregulated. Using gene ontology search terms, the highest scoring pathways that were altered at the lower concentrations were associated with inflammation and mitogen-activated protein kinase signaling (which controls inflammation). At the higher concentrations, stress response and apoptosis pathways were altered. Inflammatory pathways were activated within the first 2 hours following exposure and returned to baseline by 24 hours. In contrast, pathways involved in apoptosis and cell proliferation were not altered until the 24-hour time point.

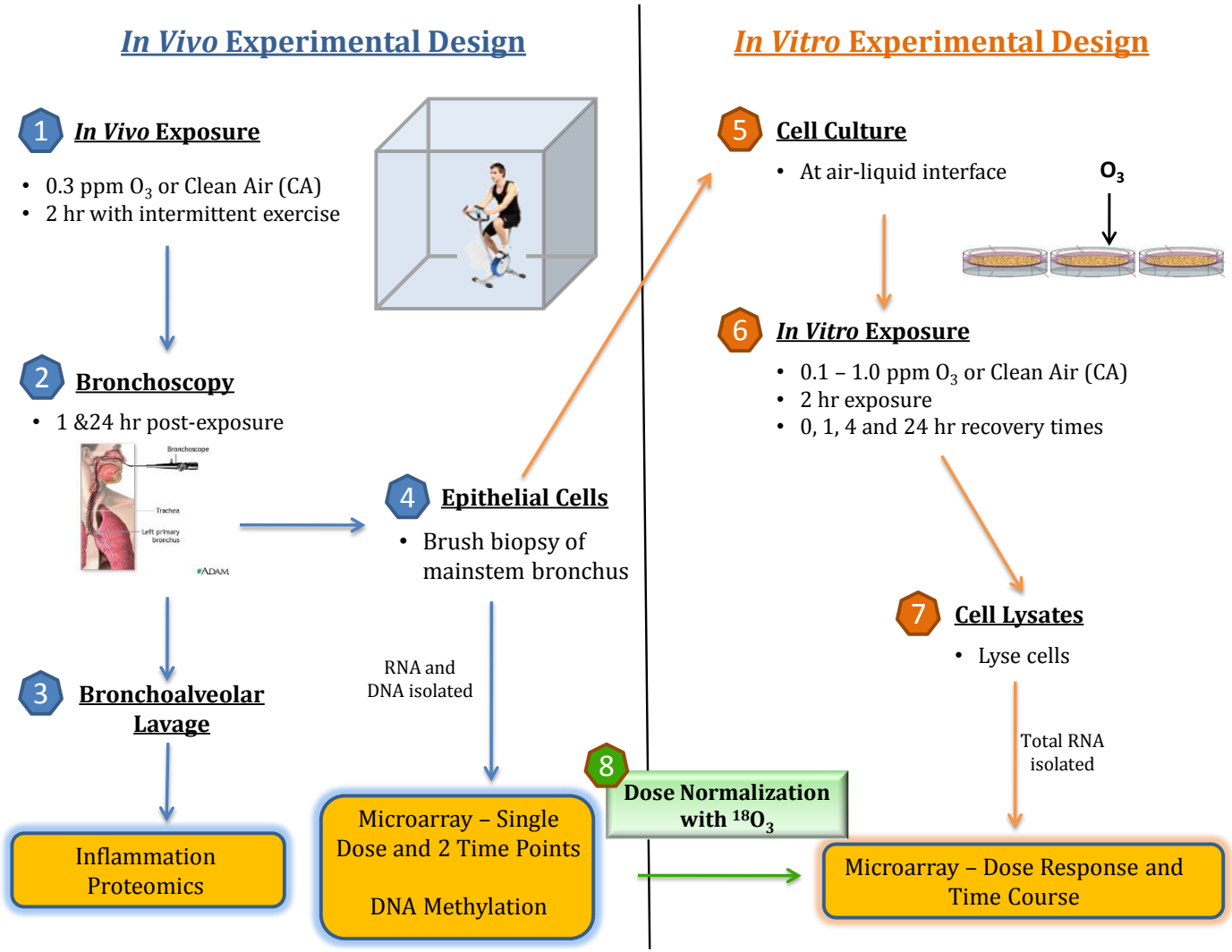


Figure 14. Ozone *In Vivo* and In *Vitro* Comparison (Devlin 2014)

This figure diagrams a study to determine whether downstream events, such as the expression of pro-inflammatory genes following *in vitro* exposure to ozone, are consistent with the expression of those genes following *in vivo* exposure. Cells from the *in vivo* exposure individuals were used in the *in vitro* studies to minimize interindividual variation.

Analysis of *in vivo* microarray data showed that more genes were activated 1 hour after exposure than 24 hours after exposure. Using Ingenuity Pathway Analysis to compare pathways altered by *in vitro* exposure with those altered by *in vivo* exposure revealed several pathways that were altered in both instances, including pathways involved in inflammation and tissue repair (cellular movement, cell-to-cell signaling and interaction, cellular growth and differentiation). These results are consistent with the published data in both animals and humans showing that ozone exposure causes cell injury, particularly of ciliated cells, and that inflammation is induced to help resolve the injury. New cells are then grown and differentiated to complete the repair of damaged tissue. Because similar pathways were altered following both *in vitro* and *in vivo* exposures, the *in vitro* response likely predicted the *in vivo* response accurately. At this time, however, these comparisons are only qualitative. Current efforts focus on developing quantitative comparisons based on the data.

3.1.2.5 Use of *In Vitro* Data to Predict Susceptibility

That individuals vary considerably in their response to ozone is well known. Controlled exposure studies in which nearly 300 healthy young individuals were exposed to multiple concentrations of ozone showed a tenfold range in lung function decrement. Individuals who returned several months later for another ozone exposure tended to retain their hierarchy on the response curve, implying that long-lived intrinsic factors could drive ozone responsiveness (McDonnell 1996). For *in vitro* toxicology to reflect *in vivo* responses accurately, cultured cells must be able to retain susceptibility factors present in live systems. Animal studies have identified several genes that are involved in ozone-induced inflammation (Bauer et al. 2010). Humans carrying the null allele of the glutathione S transferase M1 gene (*GSTM1*), a phase 2 antioxidant gene, have been shown to have increased ozone-induced pulmonary inflammation compared with individuals carrying the wild type allele (Wu et al. 2012). Cultured lung epithelial cells obtained from individuals carrying the *GSTM1* null allele have been shown to be more responsive to ozone than cells obtained from individuals carrying the wild-type *GSTM1* allele (Figure 15). Wu et al. (2011) indicated that at least some of these susceptibility factors can be studied using *in vitro* approaches.

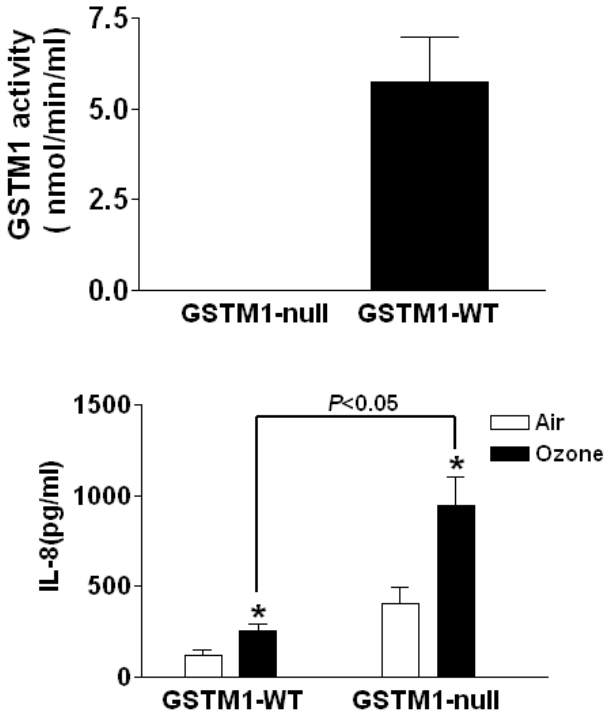


Figure 15. *GSTM1* Modulation from Bronchial Epithelial Cells Exposed to Ozone.
Top panel: *GSMT1* activity cells derived from null or wildtype individuals. Bottom panel: IL-8 activity following air and ozone exposure in null and wildtype individuals.

3.1.2.6 Risk Assessment Implications Based on the Ozone Prototype: Use of New Data

Hazard Identification

In controlled human and *in vitro* experiments, the likely causal nature of gene/pathway changes in the induction of lung inflammation in response to ozone has been described (McCullough et al. in press). Similar pathways are induced in epithelial cells exposed to ozone *in vitro* and in epithelial cells removed from humans after ozone exposure. These pathways include those involved in inflammation and tissue repair. The pathway information, coupled with *in vitro* data about ozone-induced changes in upstream signaling pathways and generation of ROS, provides a better characterization of the AOP network and increases confidence in predictions based on *in vitro* data for downstream i*n vivo* pathophysiological changes. The *in vitro* airway epithelial cell model used here might be further developed and validated for use in predicting the potential for similar inhaled chemicals (e.g., those that cause oxidative stress) to induce *in vivo* inflammation. A high-throughput screening assay based on this cell model is currently under development, and could greatly improve our ability to provide rapid hazard identification for a much larger domain of chemicals.

Exposure-Dose-Response Assessment

The analysis of transcriptional changes across a range of doses and time points for the *in vitro* experiments was not feasible for the human *in vivo* experiments reported here because of the time required to perform controlled human exposure studies. Differences in gene expression were observed that were both dose and time dependent, indicating the importance of characterizing these variables when extrapolating from *in vitro* to *in vivo* effects.

Cumulative Risk Assessment

Many air pollutants appear to induce cardiopulmonary inflammation, which likely plays a role in risks for asthma, emphysema, and pulmonary fibrosis. Developing the ozone AOP network (also called MOA) provided useful insights into the molecular network leading to lung inflammation and disease and into how air pollutant exposures cause inflammation. Using *in vivo* approaches is not feasible to evaluate cumulative risks from exposures to the numerous potential mixtures of inhaled toxicants to which people are exposed. High-throughput *in vitro* approaches, however, could screen hundreds of different combinations for potential disruption to the AOP and identify a small number of especially relevant mixture combinations for further *in vivo* approaches.

Variability and Susceptibility in Human Response

Not all individuals are equally responsive to toxicants; some are much more responsive because of age, gender, race, disease, lifestyle (e.g., smoking, obesity), or genetic/epigenetic factors). If *in vitro* assays are to predict *in vivo* responses, they must account for differential responsiveness. The experiments with *GSTM1* null and wildtype individuals suggest that at least some of these factors can be modeled using *in vitro* approaches. Asthmatics have an enhanced inflammatory response to ozone (Bosson et al. 2003; Peden et al. 1997). A recent study shows airway epithelial cells obtained

from asthmatics appear to retain an asthma phenotype[19] in culture and are more responsive to pollutants than cells obtained from nonasthmatics (Duncan et al. 2012). These results suggest that both genetic and disease susceptibility might be modeled using *in vitro* approaches

Some investigators might have difficulty obtaining primary airway epithelial cells for *in vitro* toxicology. Recent advances have been made, however, in the use of inducible pluripotent stem cells derived from adult skin or blood cells to generate cells of different phenotypes, including lung epithelial cells (Wong and Rossant 2013). Such advances offer the promise of obtaining unlimited cells from large numbers of individuals with different types of pollutant susceptibility.

Ozone is one of the few pollutants for which an extensive animal and human health effects database is available. Coupled with *in vitro* pathway data, this prototype pollutant can be used to illustrate how a systems biology approach can be used to estimate low-exposure responses in humans, to extrapolate between *in vivo* and *in vitro* human data, and perhaps to account for various susceptibility factors. Additionally, this prototype illustrates how genomics data can be used in the risk assessment of inhaled pollutants.

Quantitative systems biology models are translational, and their development is data driven. Model structure and dynamics are parameterized using data on basic biology, how that biology is perturbed by toxicants, and how and when adaptive or adverse responses develop. Systems biology models that are sufficiently well developed and well validated can be used to predict dose-response and time-course behaviors for pathway perturbations, adaptive responses, and adverse health effects. The accuracy and usefulness of those predictions, however, greatly depend on the extent and quality of the data used as inputs, and on the technical quality of the model itself. Time-course and dose-response pathway data from *in vitro* exposure studies paired with pathway data from *in vivo* exposure studies are needed to model ozone toxicity pathways responsible for downstream pathophysiological changes such as inflammation. Such data provide the detailed information needed to develop models that can predict key events like those illustrated in Figure 9.

Challenges in the Development of HT Assay Development

Using information based on *in vitro* pathways to predict human *in vivo* responses to toxicants for risk assessment purposes presents certain challenges. A major hurdle relates to extrapolation from *in vitro* to *in vivo* effects. Many *in vitro* approaches use animal cells or transformed cell lines derived from humans, which might not accurately reflect cell interactions or events in the pathway for human *in vivo* effects. The data shown above demonstrate that the response of a primary cell to a toxicant can significantly differ from the response of a transformed or immortalized cell line. A parent toxicant might also need to be biologically transformed into a more active form by cells that are not represented in the *in vitro* system (e.g., liver cells) before interacting with the target cells represented in the assay. In the lung, however, epithelial cells that line the human airways are the first and primary targets of inhaled toxicants and are believed to be the cells that initiate lung

[19]**Phenotypes** are the observable physical and biochemical characteristics of gene expressions; the clinical presentation of an individual with a particular genotype.

inflammation. Biomarkers produced by cultured cells exposed to air pollutants are also found in the lung following *in vivo* exposure to the same pollutant (Selgrade et al. 1995). This ability to show concordance between pathway perturbations *in vitro* following exposure and pathway perturbations *in vivo* is critical, and a major advantage of the model lung system discussed here.

A second challenge associated with *in vitro* approaches is ensuring that the dose of toxicant delivered *in vitro* to cultured cells can be extrapolated to estimate a comparable dose to target cells from an *in vivo* exposure. In most *in vitro* studies, cultured cells are exposed to toxicant levels that might be orders of magnitude greater than would be experienced *in vivo.* This disparity raises the uncertainty as to whether the same biological pathways are adversely affected in both situations. In the ozone prototype, $^{18}O_2$ was used to deliver an *in vitro* dose relevant to the concentration used in the *in vivo* studies. This approach has been used previously to normalize the dose of ozone delivered to rats and humans (Hatch et al. 1994). The approach increases confidence in estimates of the animal-to-human extrapolation of target tissue doses, that is, that target tissue doses in rats exposed to 2.0 ppm ozone are comparable to target tissue doses in humans exposed to 0.4 ppm ozone. This same approach can be used to normalize the dose of ozone delivered to cultured cells and humans.

In summary, the results of the ozone prototype model support the use of toxicogenomic data to identify relevant molecular pathways and network disruptions associated with adverse outcomes following exposure to a specific toxicant. Anchoring molecular patterns to adverse health effects requires considerable high-quality data and systems biology knowledge. In the case of ozone, the new approaches provide sufficient knowledge about the pathways in the network for air pollutant-induced inflammation to develop high-throughput screening (HTS) assays that can screen chemicals for potential *in vivo* effects.

3.1.3 Tobacco Smoke-, PAH-, and B[a]P-induced Cancer

PAHs are a group of over 100 different chemicals that share the feature of being neutral, nonpolar hydrocarbons with structures composed of different numbers of fused aromatic rings (rings of alternating double and single carbon atom bonds). They are formed during the incomplete burning of carbon-containing materials like coal, oil, and gas. They are also found in other organic substances that have incomplete combustion (due to insufficient oxygen or other factors) such as tobacco smoke and some foods (charbroiled meat). PAHs exist in the environment almost exclusively as complex mixtures, and are a major component of urban air pollution. Solubility in water is low, but PAHs can contaminate drinking water (e.g., from oil spills) and be taken up in the food chain.

PAHs generally have a low degree of acute toxicity in humans. The most significant adverse effect from chronic exposure to PAHs is cancer. Several PAH-containing complex mixtures such as coke oven emissions, diesel exhaust, and tobacco smoke are considered carcinogenic in humans. Most of the experimental data come from animal studies, however, increased incidences of lung, skin, and bladder cancers have been associated with occupational exposures to PAHs. Data for other sites are much less persuasive (ATSDR 1995; IARC 2010).

Ascribing observed health effects in epidemiological studies to specific PAHs is difficult because most exposures to PAHs are in mixtures. Many individual PAHs have sufficient experimental evidence to be considered carcinogenic in humans, and these vary in potency (IARC 2010). Given the universe of PAHs and potential PAH-containing mixtures, testing all of the varieties and potential mixtures with traditional approaches is not feasible. Compounding the challenge is the fact that many PAHs are only slightly mutagenic or even nonmutagenic *in vitro*, but their metabolites or derivatives can be potent mutagens. Accounting for metabolic capability and variability in assessing risk to PAHs is therefore also important.

Newer methods and data to assist in this effort include pathway mining and computational models. The prototypes presented in this section demonstrate pathway mining techniques to assist risk assessor or assessment teams in systematically searching existing toxicogenomic data, re-analyzing the data, and interpreting the data for use in risk assessment. An example is provided of pathway mining to identify whether human transcriptomics data from cigarette smoke (a complex mixture of chemicals that includes PAHs) could be associated with lung cancer. The point is not to demonstrate that cigarette smoke is causally associated with lung cancer, but rather to demonstrate how associations between chemical exposures and defined diseases could be identified for chemicals with limited or no traditional data but with molecular data using this methodology.

A second prototype is presented for B[a]P. B[a]P is one of the most studied PAHs, and the available data are sufficient to develop hypotheses about molecular pathways, and to simulate the pathway dynamics with a computational model, in this case, a Boolean model. Model runs then can be used to test hypotheses for pathway dynamics or to estimate how potent a PAH might be based on how much it perturbs key nodes in the simulated network. Both approaches could provide new tools for evaluation of complex mixtures.

3.1.3.1 Systems Biology Approach to the Assessment of Tobacco Smoke (a Complex Mixture of PAHs and Other Chemicals)

A prototype was developed that compared the toxicogenomic data from smokers and nonsmokers to determine if smokers have more similar gene expression changes than nonsmokers to the gene expression changes seen in lung cancer phenotypes. The prototype was designed to address two questions: (1) Can toxicogenomics data be used to support causal inference about a hazard? and (2) Can toxicogenomics data be used to test an MOA hypothesis?

The null hypothesis is that smokers do not have more similar gene expression changes in the hypothesized pathways for the MOA for lung cancer compared to nonsmokers. Toxicogenomics data from smokers is used to test the alternative hypothesis that smokers do have more similar gene expression changes. The hypothesized MOA is based on existing disease pathway gene expression data (i.e., pathways that are perturbed in support of a disease state) for lung cancer. The general scheme for the hypothesized toxicity pathways leading to cancer from tobacco smoke is illustrated in Figure 16. Loss of normal growth control is generally thought to result from increased inflammation and DNA damage.

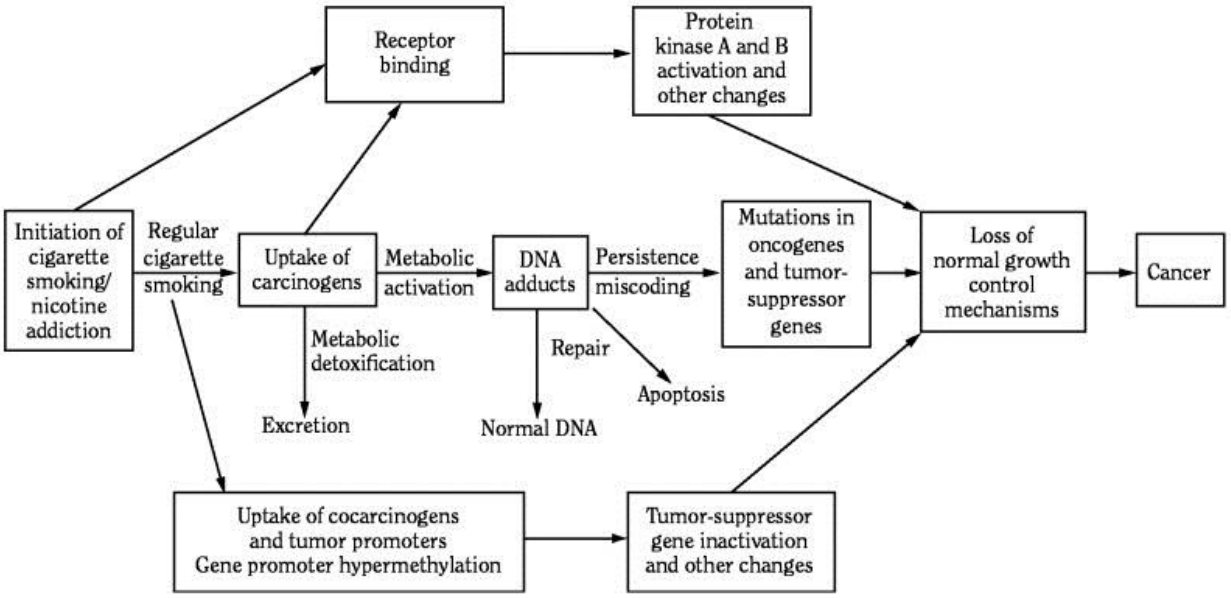


Figure 16. Adverse Outcome Pathway (AOP) for Cigarette Smoking-induced Cancer (2010).

If the null hypothesis (i.e., smokers have more similar gene expression changes with the expression changes in the disease pathways than nonsmokers) is rejected, this information is still inadequate to assert the MOA; however, it does provide limited support. In combination with additional MOA-focused studies, this information will help inform the MOA.

Pathway Mining Method To Test for the Association Between Smoking and Cancer

We first determined that human transcriptomics data from cigarette smoke (a complex mixture of PAHs and other chemicals) could be associated with lung cancer.[20] To resolve key components of the tobacco smoke-cancer AOP network, the Gene Expression Omnibus (GEO) and ArrayExpress gene expression data repositories were systematically searched for existing data using the following key word queries and the following results:

- "Cigarette smoke lung cancer" – 47 entries
- "Cigarette smoke" – 444 entries
- "Lung cancer" – 5,046 entries
- "Small cell lung carcinoma" – 166 entries

[20]An overview of this work was presented at the National Academy of Science's Emerging Science for Environmental Health Decisions Meeting on Mixtures and Cumulative Risk Assessment (Burgoon 2011).

For use in risk assessment, access to the raw data is needed (i.e., for transparency), and the data must be peer-reviewed. If the raw data were not available, the study results were excluded from further consideration. To determine whether a study was peer reviewed, the study listed in GEO or ArrayExpress needed to have an accession number.[21] If GEO or ArrayExpress did not list the study, we performed PubMed and Google searches using the author names to identify whether the GEO or ArrayExpress accessions were listed in the paper(s). Studies that were not associated with a peer-reviewed paper were excluded from further consideration.

Two studies from GEO met the above criteria: GSE10072 and GSE5060. Both are studies of human lung tissues. GSE10072 provided data on smokers and nonsmokers who were positive or negative for adenocarcinomas; this was used to derive the disease pathway. GSE5060 provided data on phenotypically normal smokers and nonsmokers; this study was used to test the hypothesis that lung cancer pathways could be detected in phenotypically normal smokers (i.e., smokers who have not yet developed lung adenocarcinomas).

Correlation-based networks (networks where probes[22] or genes are connected based on expression similarity) were built from the lung tumor data and the smoking data. Four expression networks were generated, one each for smokers, nonsmokers, lung tumors, and normal phenotype. Networks specific for smokers and lung tumors were identified by first subtracting out the network for nonsmokers from the smoker network, and subtracting out the network for normal phenotype from the lung tumor network. This resulted in two "difference networks," one for smokers and one for lung tumors. The lung tumor and smoker networks then were intersected, resulting in a subnetwork of only those probes that are connected to each other in both networks. In other words, the probes in the intersected network are connected to the same probes in both the lung tumor and smoker networks.

The data mining and subnetwork (i.e., intersected network) approach discussed above identified probes associated with both exposure and disease, and provides more insight into the MOA for a disease than the typical toxicogenomic study results. For example, in the intersected network, several communities or regions are more highly connected to each other than to other nodes in the network. The first community consists of three probes representing the pRB (retinoblastoma) gene. The *pRB* gene is a G1/S-phase cell-cycle transition checkpoint regulator. The three pRB probes are connected to the *Pak3* gene. The *Pak3* protein is activated by p21, Cdc42, and RAC1, all of which are involved in cell motility and proliferation associated with cell cycle progression. One of the pRB probes is also connected to a probe for the CDX1 gene, which is normally expressed in the intestine. In lung tumors, CDX1 is hypothesized to play a role in the aberrant expression of MUC6 (mucin). MUC6 co-expression with MUC2 is associated with poor prognosis in small adenocarcinomas. One of the other pRB probes is connected to PIGO (phosphatidylinositol glycan anchor biosynthesis, class O) and CYP4A11. These two genes are connected metabolically in glycan and phospholipid

[21]An accession number is a unique identifier assigned to a particular genome or protein sequence to uniquely identify it in a database.

[22]**Probe** – a term to describe a reagent used to make a single measurement in a gene expression experiment.

synthesis/metabolism, and are involved with metabolic production of 20-HETE (20-hydroxyeicosatetraenoic acid), which is associated with cancer cell proliferation (Alexanian and Sorokin 2013).

Advantages of the Data Mining and Intersecting Network Approach

Toxicogenomic studies, as they are typically performed today, identify potential pathways and modes of action. They support hypothesis generation, not hypothesis testing, and thus are inadequate for use in supporting a conclusion or as the basis for a decision in a risk assessment. Because using the same data to derive and test a hypothesis is generally poor practice, risk assessments relying on these typical toxicogenomics data sets to inform MOA arguments would require data to develop a hypothetical MOA, as well as data to test that hypothesis. The pathway mining and subnetwork (i.e., intersected network) approach identifies associations between chemical exposure and disease networks, and thus provides more insight into the likely MOA. Although the resulting data remain insufficient to definitively verify a chemical's MOA, this pathway mining approach makes better use of the wealth of information available in the disease and pharmaceutical research literature, provides a more targeted evaluation of the MOA, and helps identify the research needed to fill data gaps. When combined with experimental data, such as pharmacological blocking of identified pathways that reduce disease risks, the generated hypotheses can be tested as shown in the benzene and ozone prototypes discussed above.

Challenges to pathway mining of the toxicogenomic data for use in risk assessment:

In developing this tobacco smoke/PAHs analysis, data access and experimental design challenges were encountered that are likely to occur in similar analyses, including:

- difficulties in obtaining the raw data required for reanalysis of transcriptomics data,
- lack of clear descriptions of the study design or analytical methods in the published article,
- use of different microarray platforms that confounded attempts to identify patterns and replicate results across multiple studies,
- different analytical methods being employed within the same platform, and
- lack of a quantitative exposure estimates, which is especially common for human studies where the exposure durations, levels, and conditions are poorly characterized.

3.1.3.2 Systems Biology Approach to the Assessment of B[a]P

In this prototype, we demonstrate how the data and a network derived from pathway mining can be used to develop a computational model for use in testing hypotheses and in evaluating chemicals for toxicity potential. The test chemical used in the prototype is B[a]P.

B[a]P is found in coal tar (from incomplete combustion). Metabolites of B[a]P are known to be mutagenic and highly carcinogenic, and B[a]P is frequently used as a positive control in carcinogenicity bioassays. Repeated B[a]P exposure has been associated with increased incidences of total tumors, tumors at the site of exposure (dietary, gavage, inhalation, intratracheal instillation, dermal and subcutaneous), and tumors at distant sites (various routes) in numerous strains and species of rodents, and several nonhuman primates (EPA 2013d).

The Data Mining Method Used To Generate the B[a]P Network

The data mining method (systematic meta-analysis) started with a search for published, peer-reviewed transcriptomics data sets using B[a]P as the test substance. The GEO and ArrayExpress databases were searched for microarray transcriptomic studies using the search terms in Table 4. The search focused on GEO and ArrayExpress as these databases contain the submitted data as raw data. Raw data are critical for the meta-analyses, especially when different analytical methods might be used to generate the study results presented in the study report.

Table 4. Search Terms and Number of Studies Retrieved from the Gene Expression Omnibus (GEO) and Array Express Microarray Repositories

Search Term	Number of Microarray Studies Retrieved
Coal tar	2
Polycyclic aromatic hydrocarbons (PAHs)	13
Diesel	11
Smoke (NOT cigarette smoke)	16
Benzo[a]pyrene (B[a]P)	53
Fuel oil	1
Cigarette smoke	63
Tobacco smoke	16

The search resulted in the identification of 26 peer-reviewed publications with 40 gene expression data sets. The adult mouse liver was chosen for the analysis here based on the number of studies available across species and tissues where B[a]P was used. Only 2 of the 26 publications reported *in vivo* transcriptomic results for the mouse liver as their primary focus. One study (study number GSE24907) was a dose-response study where five male Muta mice (a LacZ transgenic mouse line) per group were gavaged with an olive oil vehicle and 25, 50, or 75 mg/kg B[a]P. The second study (study number GSE18789) was a time-course study where 27- to 30-day-old B6C3F1 mice were gavaged with 150 mg/kg B[a]P for 3 days and sacrificed at 4 or 24 hours after the final dose.

The Systematic Omics Analysis Review (SOAR) Tool was used to document and initially evaluate both studies for quality. SOAR consists of 35 objective questions that help users determine if a study contains data of sufficient quality for use in a risk assessment context. SOAR was developed by toxicology and toxicogenomics experts, and based, in large part, on existing and published data standards such as the Minimum Information About a Microarray Experiment (MIAME) standard. Both studies (GSE24907 and GSE18789) met the SOAR screening threshold. In the follow-up, in-depth scientific review, both studies were also found to be of sufficient quality for use.

That lists of differentially expressed genes reported in the peer-reviewed literature are not reproducible even across similar studies is generally known among researchers (Chuang et al. 2007; Ein-Dor et al. 2005; Fortunel et al. 2003; Lossos et al. 2004; Shi et al. 2008). In one published

example, three different studies designed to identify "stemness" genes[23] yielded 230, 283, and 385 active genes, respectively, yet the overlap for same genes expressed among the three studies was only 1 gene (Fortunel et al. 2003). A pathway-based meta-analysis uses a standardized analysis and ranking according to fold change, or a more formal meta-analyses of the raw data (Chuang et al. 2007; Ramasamy et al. 2008; Shi et al. 2008). A pathway-based meta-analysis approach is considered to be more reproducible than published differentially expressed genes results.

The B[a]P Network

Both identified studies were reanalyzed independently at the feature level[24] using the same pre-processing, normalization, and analytical methods. GeneGo Metacore was used to identify pathways representing a large number of genes from both data sets. A consensus systems model was synthesized based on the results from GeneGo Metacore (2013f) and Burgoon (2011) (Figure 17 and Table 5). The core processes represented in the network are the induction of DNA adducts, mediation of p53 (a tumor suppressor gene) signaling, alteration of translesion synthesis,[25] and regulation of the G1/S-phase transition and cell cycle. Based on the network interactions, DNA adducts are believed to be formed by reactive B[a]P metabolites generated via induction of cytochrome P450 (CYP) enzymes, secondary to B[a]P activation of the aryl hydrocarbon receptor (AhR). Others have shown AhR-independent DNA adduct formation, raising questions about other non-CYP1A1- and CYP1A2-mediated B[a]P metabolism and adduct formation (Kondraganti et al. 2003; Sagredo et al. 2006).

The consensus model in Figure 17 conceptually describes the events that might occur when B[a]P enters the cell. Briefly, B[a]P binds to AhR, leading to upregulation of xenobiotic metabolizing enzymes and Nrf2, which might lead to additional B[a]P metabolism to epoxides, and increased oxidative stress. B[a]P-mediated genotoxicity, evidenced by DNA adducts, occurs and activates p53. Although Nrf2 is upregulated transcriptionally, p53 is expected to interfere with Nrf2 signaling, ensuring a pro-oxidant environment, which could perpetuate further DNA adduct formation. Upregulation of p21 (Cdkn1a) and MDM2 are most likely a result of p53. Upregulation of ubiquitin, while in the presence of p53-mediated MDM2 upregulation, is expected to destabilize p53.

---

[23]Stemness genes are those genes that are hypothesized to confer stem cell characteristics.

[24]A common misconception about microarrays is that they measure gene expression at the level of a gene. In reality, microarrays measure only a portion of a gene, typically anywhere from 20 to 100 nucleotide bases. This portion of the gene that is actually measured is called a "feature." Typically, only one feature exists per gene on a microarray. Some genes are represented more than once on a microarray, however, complicating downstream analyses (e.g., deciding how much a gene is expressed when the two features representing different parts of the same gene yield different numbers). Features also could be believed to map to a specific gene at one time, and the feature is later discovered to map to a completely different gene (this happens more frequently with lesser known or studied genes and lesser known or studied organisms where the genome might not be available). Thus, the gene associated with a feature can change over time, and most analysts will remap their feature sequences against the genome periodically to ensure they have the latest annotation. This might result in reproducibility issues when comparing to studies performed at different times. Generally, when interpreting gene expression, analysts prefer to operate at the feature level for all analyses.

[25]Translesion synthesis is a mechanism that the cell uses to continue DNA replication/synthesis in the presence of a DNA lesion (e.g., DNA adduct).

Destabilization of p53, in the presence of proliferating cell nuclear antigen (PCNA), is expected to allow translesion synthesis, resulting in mutations and adducts that persist through DNA synthesis. Upregulation of Cyclin D could be sufficient to overcome p21 inhibitory competition, especially as p53 levels decrease, allowing for G1/S-phase transition to occur. G1/S-phase transition, combined with translesion synthesis, would lead to propagation of mutations and DNA adducts into daughter cells. This could continue as a feed-forward loop until p53 signaling is reinitiated.

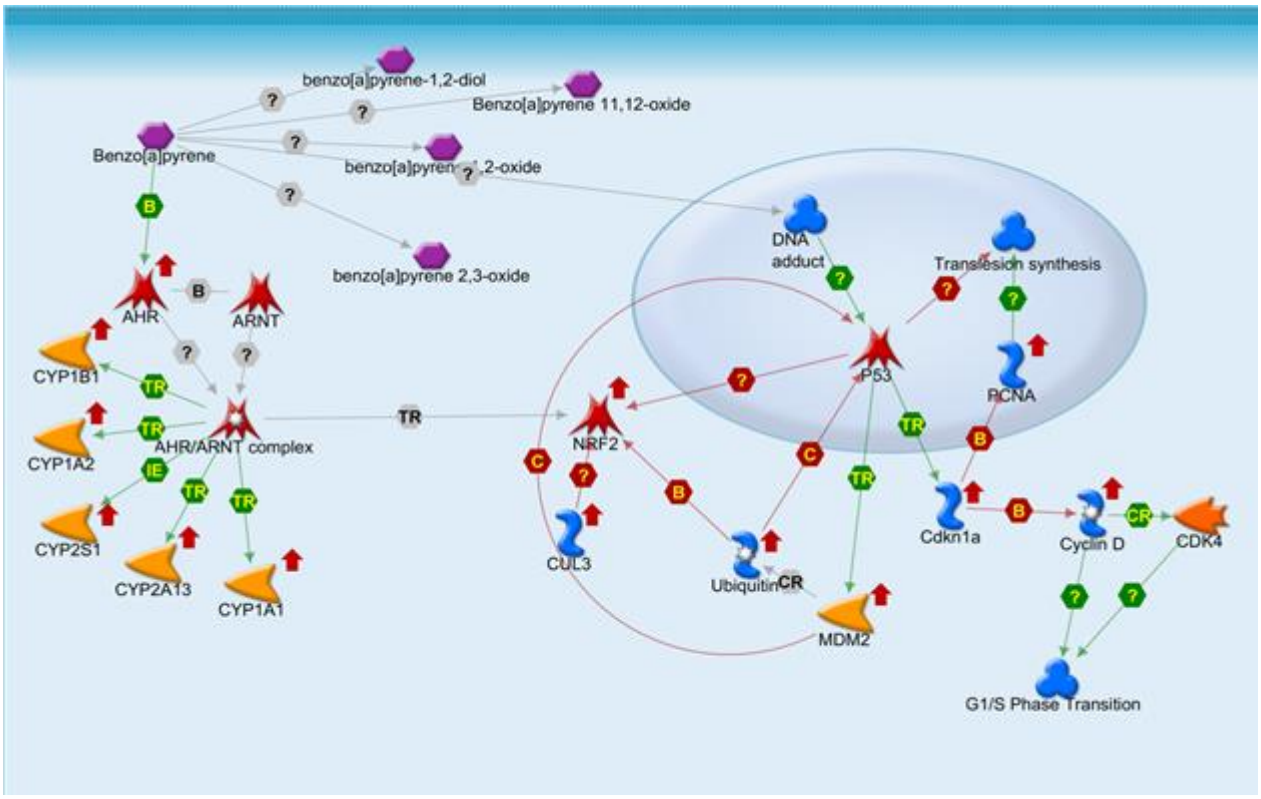


Figure 17. Consensus Outcome Pathway.

This consensus pathway was synthesized by combining multiple pathway diagrams identified through analysis of the two data sets using GeneGo Metacore. The nodes (proteins or outcomes) are connected by lines. The green lines represent activation, while the red lines represent inhibition or repression. The thick red arrows near proteins represent increases in gene expression. Hexagons with a "TR" represent a transcriptional regulation relationship (e.g., AhR/ARNT complex transcriptionally regulates CYP1A1), those with a "B" represent a case where one protein binds to another, "IE" represents an indirect regulation of gene expression, "CR" represents a case where proteins form a complex, while a question mark represents an unspecified relationship or interaction. The GeneGo map legend can be found at: http://pathwaymaps.com/pdf/MC_legend.pdf.

Table 5. Altered Genes/Functions and Their Relationship to Cancer (in this Model)

Altered Gene or Function	Relationship to Cancer in this Model
AhR/ARNT Complex	AhR regulated expression of several CYPs, including CYP1A1 and CYP1A2
CYPs (e.g., CYP1A1, CYP1A2)	Upregulation leads to production of oxidative radicals and B[a]P metabolites
NRF2	Regulates the expression of oxidative stress-protective genes
Ubiquitin	Protein that tags other proteins for destruction
CUL3	Regulates the inhibition of NRF2 signaling with ubiquitin
p53	Stops cell cycle by preventing G1/S-phase transition; activated by DNA damage
MDM2	Regulates p53 through negative feedback mechanism with ubiquitin
Cdkn1a/p21	Upregulated by p53 activation; inhibits Cyclin D activation and prevents G1/S-phase transition
Cyclin D	Activates G1/S-phase transition, works with CDK4
CDK4	Activates G1/S-phase transition, works with Cyclin D
G1/S-Phase Transition	Starts cell cycle progression by allowing for DNA synthesis
Translesion Synthesis	DNA damage tolerance mechanism; allows DNA replication fork to progress beyond DNA damage sites
DNA Adduct	A piece of DNA covalently bound to a chemical that can modify expression of DNA

Developing a Computational Model (Boolean Network Model) Based on the AOP Network

Based on the gene expression changes and activating DNA adduct formation, a Boolean Network (BN) systems model was developed (Figures 18–20) that can be used to predict the activation of cell cycle progression with translesion synthesis (Figure 21). In a BN model, system dynamics are simulated by a series of connected nodes where each node represents a gene/protein, and the connections between nodes (edges) represent some type of action/inhibition relationship. The connections are directed. For example, p21 inhibits Cdk4, so the arrow originates at p21 and terminates at Cdk4. Some actual relationships are not as simple, for example, Cyclin D interacts with Cdk4 to activate G1/S-phase transition. To also represent the p21 relationship with Cdk4, it is best to represent the Cyclin D action on G1/S-phase transition in the BN model as a positive interaction between Cyclin D and Cdk4.

Each node is in either an on (1) or off (0) state. The Boolean Network cycles through different overall system states, based on changes in the state of each node in relationship to the other nodes over time. To test a hypothesized outcome (e.g., that cell cycle progression and translesion synthesis will be sustained once initiated), the BN model was simplified to represent just the DNA adduct/cellular proliferation processes. Model runs were then conducted. Of interest here is the occurrence of stable states or attractors, that is, cycles of states that recur and self-perpetuate. As the BN model runs progress, states that become attractors are called the "basin." The Boolean Network in Figure 18 has a single state attractor defined as a cell-cycle progression state with

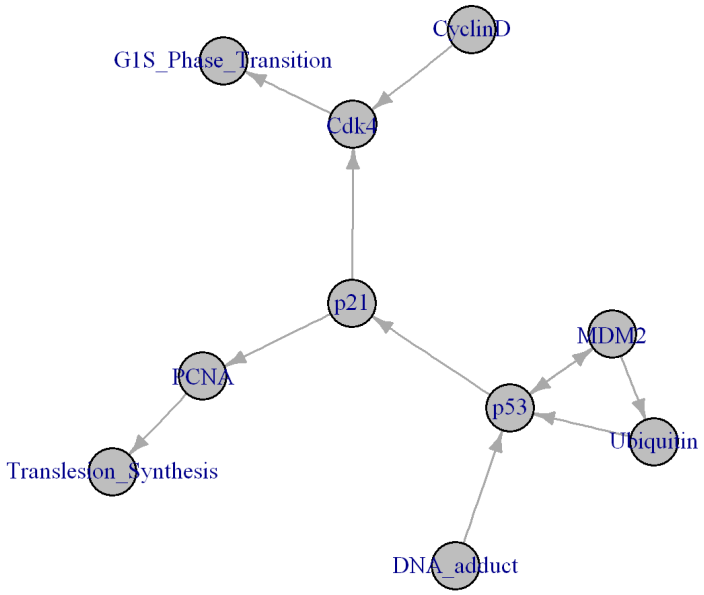


Figure 18. Liver Carcinogenesis Boolean Network (BN) Systems Model.
The nodes represent proteins, and the lines are directional connections meaning activation or inhibition (activation and inhibition are not treated differently in the graphical depiction of the model). For instance, the arrow from proliferating cell nuclear antigen (PCNA) to translesion synthesis means that PCNA activates translesion synthesis. The two major outcomes in this model are translesion synthesis and G1/S-phase transition. The major external input is deoxyribonucleic acid (DNA) adduct formation. DNA adducts cause structural damage to the DNA, which could become or lead to mutations and ultimately tumorigenesis and cancer.

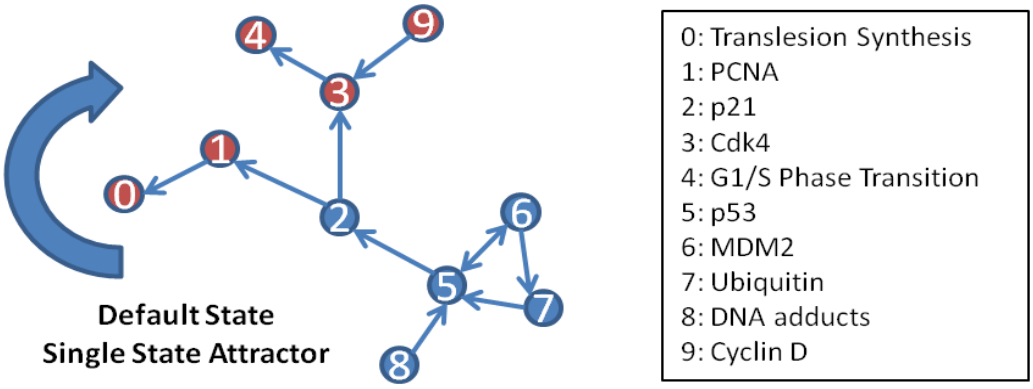


Figure 19. Default State, Single State Attractor.
The systems model falls into a default state, single state attractor system. This is the same as the network represented in Figure 18. The names have been replaced by numbers, which are noted in the figure legend. Red nodes are those that are activated. Blue nodes are inactivated. The system here has not been perturbed by external forces. Of particular interest is that the "default" state for the system is one where the cell is actively proliferating and undergoing translesion synthesis.

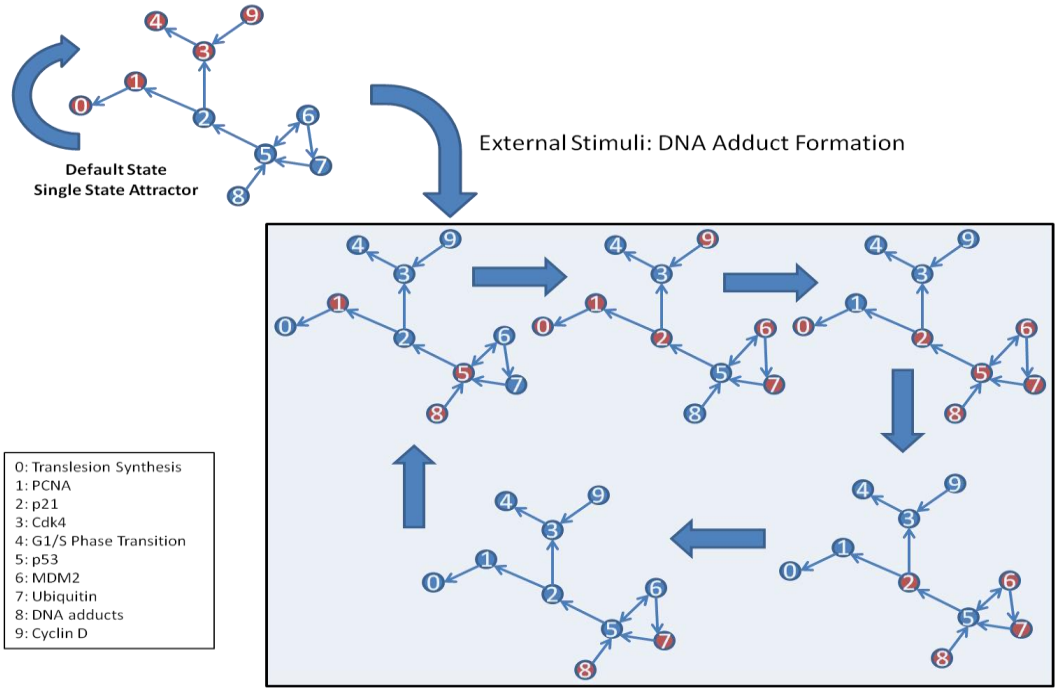


Figure 20. Deoxyribonucleic Acid (DNA) Adduct Attractor System.
When the systems model is perturbed through an external stimulus (DNA adduct formation), it transitions from the default stable starting state and moves to a new attractor (depicted in the inset). Once the system moves out of the basin for the default state attractor, it cannot return to that state without another significant stimulus. This multistability (the fact that a system can have multiple stable attractor states) is a characteristic of complex systems. Starting at the upper left of the inset, proliferating cell nuclear antigen (PCNA) is activated, DNA adducts are activated, and p53 is activated. This leads to translesion synthesis and activation of p21, MDM2, and ubiquitin. Although Cyclin D gets activated, there is no activation of G1/S-phase transition. The system then transitions to a state where translesion synthesis is primed and ready to go. If G1/S-phase transition were to occur, p53 is activated, along with DNA adduct formation, MDM2, and ubiquitin. The next system state has continued p21 activation, loss of p53 activity presumably through ubiquitin and MDM2 activation in the prior system state, and DNA adduct formation. The system then transitions to only DNA adduct formation and ubiquitin activation, followed by restarting of the cycle.

translesion synthesis turned on, here designated as $state_{TL}$, and presented in Figure 19. If the cell were to enter this $state_{TL}$, it would be expected to self-perpetuate until a stimulus altered the system in a way that increased the frequency of other states. Importantly, the current BN model does not predict that all cells will enter $state_{TL}$ or that $state_{TL}$ is the default. Rather, model runs simply indicate that if $state_{TL}$ were entered, the cell would remain in $state_{TL}$ until a stimulus occurs sufficient to change the dynamic and transition the system to a different state. Such stimuli might include changes in gene expression, alterations of metabolic status, or a change in overall energy level.

The current BN model predicts that, with DNA adducts alone, the cell will enter into a five-state attractor (Figure 20). In this cycle, the cell is not predicted to enter into G1/S-phase transition, which one would expect because p53 should effectively shut down that pathway. Translesion synthesis is predicted to occur in this five-state attractor cycle.

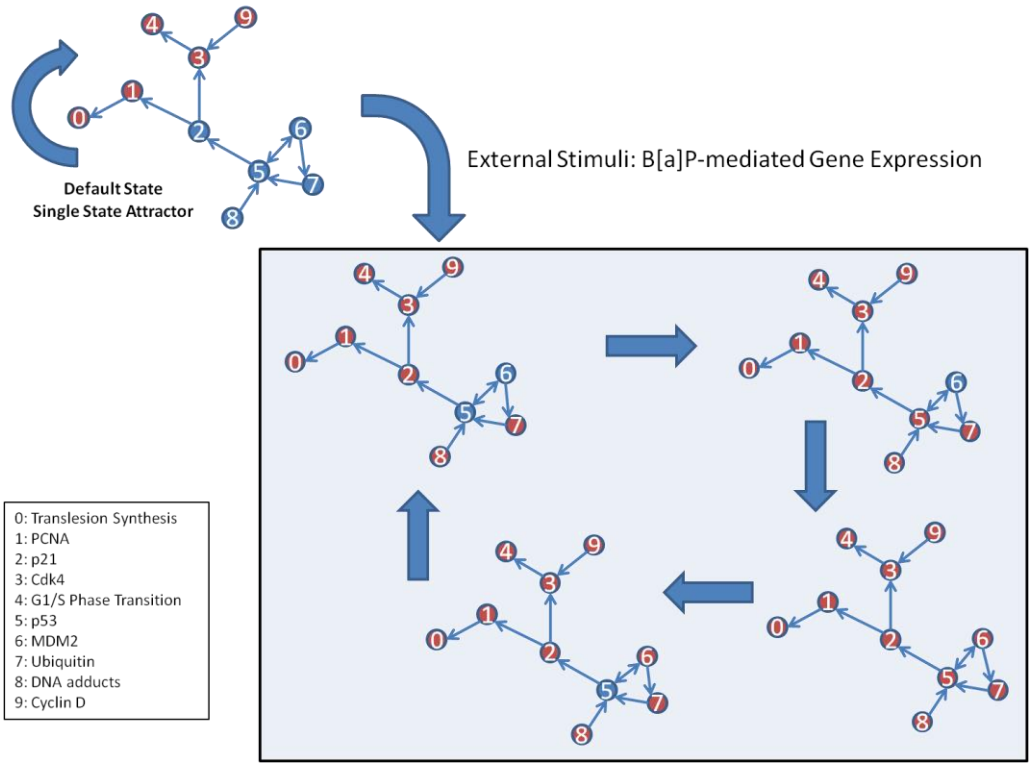


Figure 21. Gene Expression Data Attractor System.
This four-system attractor is based on the gene expression data observed in both studies. This attractor system is notable as it shows deoxyribonucleic acid (DNA) adduct formation, translesion synthesis, and G1/S-phase transition occurring in all system states. This model predicts that DNA adducts and potential mutations are being passed forward to daughter cells through translesion synthesis as the cell cycle progresses at these doses and times in the mouse liver. This suggests that B[a]P at these doses and experimental time-points post exposure in the mouse liver could be an initiator and promoter of tumorigenesis. This adverse outcome pathway (AOP) might ultimately result in carcinogenesis.

Given that the data and model representation of the system dynamics are reasonable and of good quality, the B[a]P BN model supports the hypothesis that high doses and acute durations like those used in the two mouse liver studies will initiate liver tumor progression through a genotoxic MOA, and that promotion occurs through a cellular proliferation MOA. The available mouse data are inadequate to simulate whether the system would be activated at low doses in the mouse. The model does, however, provide a hypothesis-testing platform for effects at lower doses, or with other species, or other PAHs, given the availability of sufficient, good-quality data. For example, transcriptomic studies with PAH mixtures, or other PAHs individually, could be conducted and analyzed to determine their impact on the proposed pathway. Gene expression data from these studies could be incorporated to elaborate the model further, and simulate additional alterations in cell behavior, compared to behaviors based solely on B[a]P exposure. The model then might be used to predict doses/exposures that lead to DNA damage, activation of translesion synthesis, or G1/S-phase transitions. The standard uncertainties when extrapolating results among species also apply to the B[a]P BN model predictions; that is, the magnitude of the dose-response effects observed in test animals might differ from what occurs in humans due to genetic or epigenetic

variability. Species differences must therefore be accounted for when interpreting model results and predictions for potential effects in humans.

3.1.3.3 Risk Assessment Implications Based on the Tobacco Smoke, PAHs and B[a]P Prototype: Use of New Data

Hazard Identification

The pathway mining and subnetwork (i.e., intersected network) approach demonstrated in the tobacco smoke/PAHs and the B[a]P analyses supports better identification of the MOA based on exposure and disease pathway associations than the typical toxicogenomic study results. The results suggest that cigarette smoke exposure activates cell cycle progression and cellular proliferation pathways. The network-based determination of pathways in the human studies demonstrated the coherence between the lung tumor and cigarette smoking molecular pathway data.

The B[a]P analysis indicates that B[a]P activates known human disease pathways associated with genotoxicity and tumor promotion/cell cycle progression. The meta-analyses of multiple animal data sets and the relatively well-understood mechanistic information provide additional confidence in outcomes indicated by the BN model and the data.

Disease-focused systems models could be developed for a larger set of complex human diseases to expand the utility of this approach going forward. Disease-state systems models would integrate metabolomics and proteomics data streams and improve our mechanistic understanding of observed dose-response relationship, in support of more rapid and accurate hazard identification screens. The genes included in a Boolean systems model could be represented in a battery of assays to be used in Tox21 screening. HTS assay batteries based on these models could be implemented using current multiplex quantitative polymerase chain reaction assay systems.

Taking the data and results from both the tobacco smoke/PAHs and the B[a]P analyses together provides sufficient support for a likely causal relationship between PAH exposure and cancer, based on the similarity of the tobacco smoke pathway activation and known cancer pathway activation in humans and the activation of known cancer pathways in the rodent B[a]P studies. Important uncertainties remain, however. Differences in tissues affected (human lung cancer and rodent liver cancer) likely depend on route of exposures. In this case, we have demonstrated coherence, as PAHs clearly act as a promoter in both the human and animal studies, triggering cell cycle progression and cellular proliferation. Observed associations in the animal studies are consistent based on the meta-analysis, but information is insufficient to demonstrate consistency of the transcriptomic pathway data in humans (i.e., only one lung tumor and one smoking data set; additional data sets needed for both). The consistency of effect in both humans and animals is not necessary for a "likely" determination based on transcriptomics, but would be helpful in advancing our understanding. Because the only data examined in this prototype are toxicogenomics data, making a strong MOA argument is not possible. To strengthen the MOA argument, additional mechanistic data consistent with the suggested MOA are needed.

Exposure-Dose-Response Assessment

The Boolean model approach supports prediction of adverse outcomes across a range of doses. The dose-response characterization in the B[a]P animal studies was sufficient to support a causal determination. Due to the lack of sufficient dose-response data, however, model simulations of the severity or incidence of adverse outcome following different PAH or B[a]P dosing was not possible (in this exercise). The B[a]P example, however, did demonstrate how the impacts of different exposure/activation scenarios could be evaluated. With sufficient dose-response data, the impact of different doses could be modeled and estimated beyond the calibration data set.

Cumulative Risk Assessment

Similar pathway-based meta-analyses could be performed on transcriptomic data for other chemicals to screen for genotoxicity and tumor promotion, prior to the observation of tumors. Adequately developed Boolean systems models could inform risk assessors of the likelihood that other PAHs or PAH mixtures share a similar MOA to the one identified for B[a]P. Models also could be developed that compare and integrate pathway-based results for multiple chemicals and nonchemical stressors, to predict outcomes from exposure to mixtures or cumulative stressors.

Variability and Susceptibility in Human Response

Variations in human genetics that alter susceptibility to tumorigenetic or carcinogenetic effects could be modeled based on data from genome-wide association studies (GWAS),[26] knock-out studies, or knock-down studies. A Boolean model would simulate and predict outcomes for susceptible subpopulations by comparing the impacts of various node or edge alterations in a network on state changes, and the sensitivity of those changes to the pathway alterations. For example, the impacts of a gene knock-out can be modeled in the Boolean systems model by consistently inactivating the node representing that protein, and monitoring how the system state dynamics are altered. As an example, SNPs are known to occur in p53, which might impact its ability to stop G1/S-phase transition. The p53 gene also has been shown to be mutated in many cancers (Vogelstein et al. 2000). Other relevant SNPs for genes or proteins can be identified using data mining approaches, and these can be incorporated into the systems model.

Population variability can be modeled using Monte Carlo simulations to estimate the risk of adverse outcomes across different genetic profiles. This would be accomplished by using the same types of models as in the human susceptibility context. Population variability would be simulated with a series of Boolean systems models, where each model represented a different subpopulation in the overall analysis at a frequency comparable to that subpopulation's occurrence in the human population (or equal to its occurrence in a hypothesized human population if performing a what-if

---

[26]**Genome-wide association study (GWAS)** is defined as an approach used in genetics research to associate specific genetic variations with particular diseases. The method involves scanning the genomes from many different people and looking for genetic markers that can be used to predict the presence of a disease. Once such genetic markers are identified, they can be used to understand how genes contribute to the disease and develop better prevention and treatment strategies.

type of scenario). For example, if 15 percent of the population is expected to experience a loss of function polymorphism, a Monte Carlo simulation would generate a 15 percent chance of choosing the Boolean systems model for the functional loss on each random draw from the population of all models.

3.1.4 Risk Assessment Implications Across the Tier 3 Prototypes

In the Tier 3 prototypes, comparisons of new and traditional data informed our understanding of the extent to which new data types can be used to make reasonable estimates of known public health risks. Benzene, ozone, and tobacco smoke/PAHs/B[a]P generate specific exposure-dependent patterns of events or AOP networks that appear causally related to specific human disease and disorder (hematotoxicity and leukemia, lung inflammation and injury, and lung cancer, respectively). These patterns are observed in humans exposed at environmental concentrations. Evidence for a causal association between these molecular patterns and specific adverse outcomes includes (1) multiple studies with similar results, (2) pharmacological interventions that reverse key events and thus block or ameliorate the adverse effect, (3) human genetic variations of unknown origin that alter the AOP network, and also alter risks of the related adverse effect, or (4) chemical and nonchemical stressors, known to cause a specific disease, and that alter the same AOP network. Additional support for the associations between specific patterns and disease outcomes comes from comparisons between human and animal data, primary cell cultures and immortal cell lines, and target and nontarget cell types and tissues. From these data, we infer the following:

- AOP networks, when sufficiently well described, can help identify hazards for data-limited chemicals based on AOP network similarities.
- AOP networks also can help (1) identify chemical and nonchemical stressors that are likely to increase risks for the same adverse effect by acting on the same AOP network (not necessarily the same key event or pathway but within the same network); and (2) better characterize susceptible (and resistant) human subpopulations based on genetic variants.
- Information that integrates diverse levels of biological organization (systems biology) is essential to link molecular events to intermediate events to adverse outcomes.
- Molecular indicators or biomarkers of exposure and effects (subsets of the AOP network) appear suitable for measuring exposure-response relationships at environmental concentrations, if sufficient sensitivity can be demonstrated.
- Although *in vitro* data can reiterate *in vivo* molecular events, differences are often observed. For the immediate future, the confidence in interpreting *in vitro* data is greater if the data are understood in the context of applicable *in vivo* data.
- Several factors add uncertainty to the use of new data types (depending on the assay protocol), including the use of cell lines versus primary cell cultures, lack of metabolic capability in certain test systems, use of target versus nontarget cell and tissue types, species differences, variability in genetic makeup, and differences in lifestage. Exposure measurement or estimate errors can be a significant source of uncertainty. All of these factors should be considered, to the extent feasible, when using new data types in risk assessment.

- Meta-analyses that integrate pathway-based data across multiple studies yield the most convincing evidence of associations among chemical exposure, AOP disruptions, and disease/disorder. Meta-analyses are generally the most appropriate method for using transcriptomics data in a risk assessment. Experimental evidence also can be a significant factor in causal determination.
- Depending on the type and level of exposure, the resulting molecular interactions might result in beneficial, adaptive, or adverse effects. The dynamic nature of these systems is complex and will require additional research to understand fully. The B[a]P prototype provides an example of dynamic modeling.
- When searching for candidate Tier 3 prototypes, one important observation was that, even among the most well-studied chemicals, very few studies reporting new data types met the data and quality criteria needed for use in risk assessment. In part, this is due to the relative newness of these data types for application in risk assessment, and the need for additional guidance on their use. Improving the utility of new data types for risk assessment will thus require explicit use of systematic data review criteria, adherence to standards of experimental and statistical practices (in data generation, analyses, and use in risk assessment), and accurate reporting of the variability and uncertainty in the data.

3.2 Tier 2: Limited-scope Assessments

Tier 2 prototypes (1) explore new types of computational analyses and short-duration *in vivo* bioassays, and (2) demonstrate some assessment approaches for limited-scope risk management decision-making (i.e., the decision context for Tier 2). Here, "limited" generally applies to chemicals with lower exposure or hazard potentials than chemicals for which a major-scope assessment is warranted, or for which the available data are so limited a major-scope assessment cannot be conducted. The amount of resources required to conduct a Tier 2 assessment is between the amounts needed for Tier 1 and Tier 3. The uncertainties in the Tier 2 assessment results are similarly ranked, more than for Tier 3 but less than for Tier 1. Intermediate testing and assessment strategies in Tier 2 aim to prioritize and quantify risk further for a potentially large number of chemicals ranked highly in Tier 1, numbers that would quickly overwhelm resources and capability to conduct traditional or Tier 3 evaluations (Thomas, R. S. et al. 2013c).

Tier 2 approaches use a systems biology approach to integrate information across different levels of biological organization—from molecules to cells to tissues to clinical outcomes—and to identify associations (or preferably causal mechanisms) between environmental exposures and outcomes, generally using relatively short-duration test methods (days to weeks). Tier 2 assessments integrate results from Tier 1 (e.g., QSAR results, HTS data) with data from advanced data mining and higher level assay systems, for example, high-content (HC) *in vitro* assays, short-term *in vivo* surrogate (e.g., zebrafish) assays, and mammalian species (rodent) assays, and computational systems biology models. Short-duration *in vivo* bioassays are relatively uncommon in risk assessments to date, but they hold great promise for providing valuable new data in the near future. Such data (see Table 6) increase confidence in the Tier 1 results, yet the approaches still can be performed more rapidly and at lower cost than a Tier 3 assessment. Tier 2 assessments also yield

Table 6. Summary of Tier 2 NexGen Approaches, Including Strengths and Weaknesses

TIER 2: LIMITED SCOPE ASSESSMENT PROTOTYPES			
	Data Mining of Existing Databases	**Alternative Species *In Vivo* Assays**	**Mammalian Short duration *In Vivo* Assays**
Approaches:	• Discovers or identifies associations among environmental exposures, study results, and human disease • Often uses meta-analyses of large existing data sets • Suggests potential adverse outcomes based on existing knowledge of other chemical-induced molecular event and disease relationships	• Experimentally measures dose-dependent, chemically induced alterations in biological functions for intact organisms using a range of specific and sensitive assays • Measures adverse outcomes that range from molecular to phenotypic changes and population effects • Uses species with shorter life spans than traditional experimental species or humans	• Experimentally measures dose-dependent, chemically induced alterations in biological functions in intact animals using a range of specific and sensitive assays • Measures molecular or cellular changes; infers potential adverse outcomes based on existing knowledge of other chemical pathway or disease relationships • Uses short-duration exposures and observation periods (hours to weeks)
Strengths:	• Significantly faster and less expensive than traditional bioassays • Uses combined data sets that include tens of thousands of humans • Integrates information across biological levels (cell, tissue, organism) and for key factors (lifestage, metabolism, species)	• Significantly faster and less expensive than traditional bioassays • Evaluates complex outcomes such as birth defects and neurobehavioral outcomes • Evaluates effects across biological levels (cell, tissue, organism) and relative to key factors (lifestage, metabolism, species)	• Significantly faster and less expensive than traditional bioassays • Includes tissue and organism integration and intact metabolism
Weaknesses:	• Observed relationships are generally associative and primarily support only hypothesis generation	• Species-to-species extrapolation is an issue • Data sample resolution for small species is often at high levels (e.g., entire organism, multiple tissues) versus likely target cells • Data on early-life exposure effects generally lacking; an exception is the embryonic zebrafish models	• Difficulties in relating events early in disease initiation process to adverse outcomes • Changes in the entire organ are often assayed rather than those in just the target cells, which can make critical changes more difficult to detect • Data on early-life exposure effects generally lacking

results for chemicals with limited data, reducing the costs and delays associated with obtaining the additional traditional data needed for a Tier 3 assessment.

Three Tier 2 limited-scope decision-making prototypes represent approaches to assessing hundreds to a few thousand chemicals. Implications for risk assessment identified by the Tier 2 prototypes are discussed at the end of this section and are integrated with other lessons learned in Section 5. The prototypes and their respective approaches are as follows:

- diabetes and obesity: knowledge mining[27] and meta-analyses of published literature,
- thyroid disruption: short duration *in vivo* assays—alternative species, and
- cancer- and noncancer-related effects: short duration, *in vivo* assays—rodent.

3.2.1 Knowledge Mining – Diabetes/Obesity

This prototype demonstrates the use of knowledge mining as a means of characterizing the associative and potentially causal relationships between disease, exposures to environmental factors, and intrinsic sources of human variability. Exploration of diabetes risks is used here as a specific example. Knowledge mining capitalizes on massive, new databases developed in recent years to organize and store data. As a condition of publication, most molecular, computational, and systems biology journals now require that the study data be submitted to specified databases. With at least 50,000 new publications each year in the field of omics[28] alone, the amount of available new data is enormous (petabytes) and growing. These databases extend across species and include substantial information on human disease. Integration and analysis of these large databases require computing and knowledge-mining techniques. Although this section focuses on knowledge mining for information on diabetes/obesity, most of the other prototypes also use knowledge-mining methods to some extent. Box 6 describes some of the challenges in resolving the type 2 diabetes molecular interaction network.

Box 6. Molecular Mechanism of Type 2 Diabetes

The development of type 2 diabetes requires impaired beta cell function. Chronic hyperglycemia induces multiple defects in beta cells. Hyperglycemia has been proposed to lead to large amounts of reactive oxygen species (ROS) in beta cells, with subsequent damage to cellular components including PDX 1. Loss of PDX 1, a critical regulator of insulin promoter activity, also has been proposed as an important mechanism leading to beta cell dysfunction. "Diabetogenic" factors include free fatty acids, tumor necrosis factor alpha, and cellular stress. These result in insulin resistance by inhibiting insulin receptor substrate 1 functions. These functions stimulate molecular mechanisms including serine/ threonine phosphorylation, interaction with suppressors of cytokine signaling, regulation of the expression, modification of the cellular localization, and degradation. Various kinases (ERK, JNK, IKKbeta, PKCzeta, PKCtheta and mTOR) are involved in this process. Although the importance of genetic factors in type 2 diabetes is little doubted, genetic analysis is difficult due to complex interaction among multiple susceptibility genes and between genetic and environmental factors. Genetic studies have therefore produced very diverse results. *Kir6.2* and *IRS*, two of the candidate genes, are known to have a central role in insulin secretion and insulin signal transmission, respectively (adapted from NCBI BioSystems Database entry; Kanehisa Laboratories 2014b).

[27]For example, the National Library of Medicine's **Gene Expression Omnibus (GEO):** a public functional genomics data repository supporting data submissions that are compliant with MIAME (Minimum Information About a Microarray Experiment). Array- and sequence-based data are accepted. Tools are provided to help users query and download experiments and curated gene expression profiles.

[28]**Omics** refers collectively to studies in genomics, proteomics, and metabolomics—research areas that collect vast amounts of molecular information on various aspects of gene expression, protein interaction, and metabolism, respectively. A PubMed search on Dec 23, 2013 for preceeding year returned the following "hits" (in parentheses) using these search terms: genome (39,571), genomic (52,861), proteome (3404), proteomic (6968), metabolome (673), and metabolomics (1778). The numbers in parentheses do not necessarily correlate with relevance, but rather illustrate the growth in new knowledge.

The risk of diabetes (and other chronic diseases) varies in the population due to genetic and environmental factors. Figure 22 presents a systems biology diagram of the complex network of interactions involved in the onset of type 2 diabetes. This network is based on meta-analysis results of multiple human studies (Kanehisa Laboratories 2014b); it is available on the National Center for Biotechnology Information (NCBI) BioSystems Database.[29]

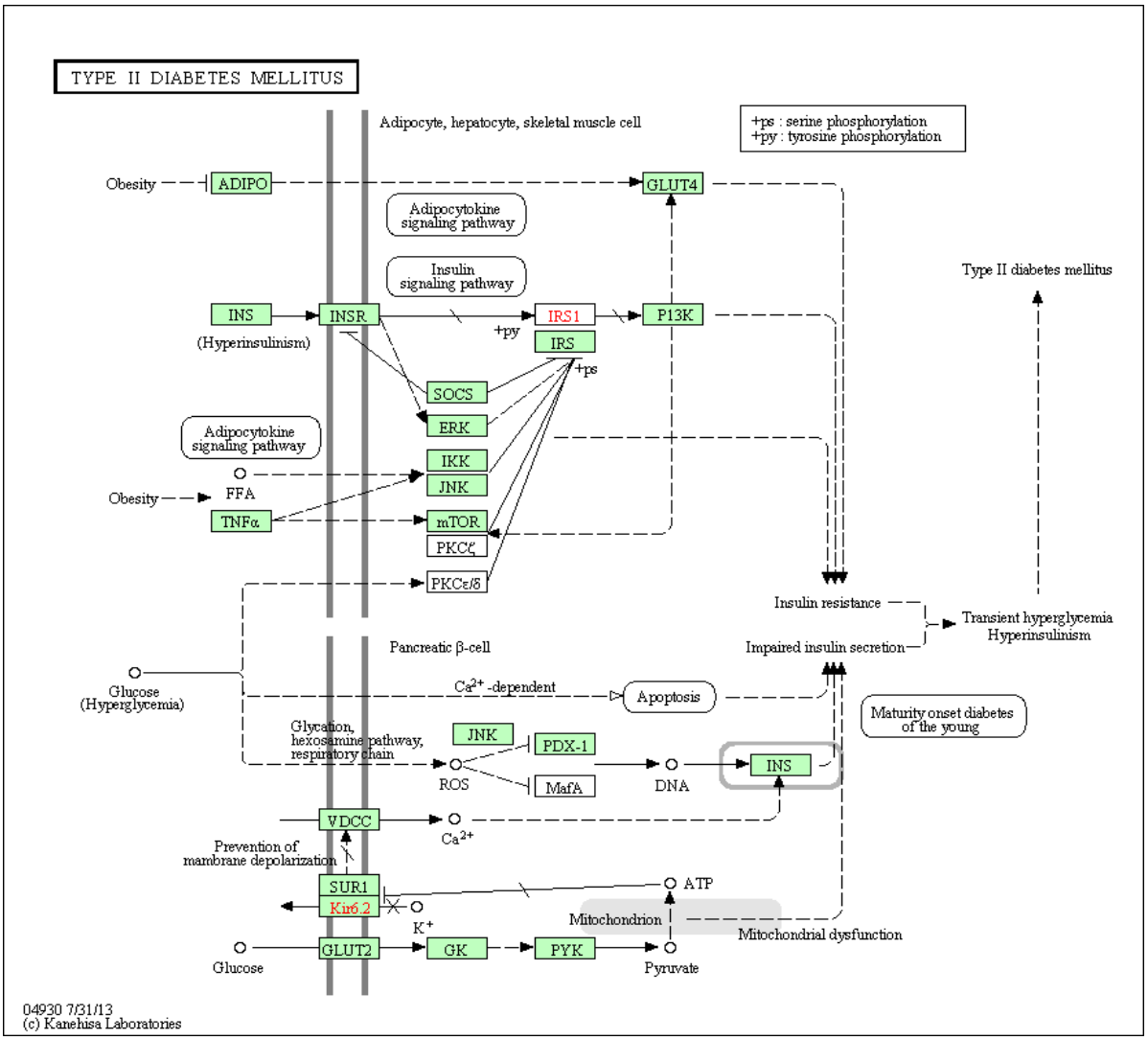


Figure 22. A Systems Biology Diagram of the Complex Network of Molecular Interactions Involved in the Onset of Type 2 Diabetes Mellitus.

This network was constructed based on the results of meta-analyses of multiple human studies, and is available on the NCBI BioSystems Database (Kanehisa Laboratories 2014b).

[29]The NCBI **BioSystems Database** was developed to (1) serve as a centralized repository of data; (2) connect biosystem records with associated literature, molecular, and chemical data throughout the Entrez system; and (3) facilitate computation on biosystems data (NCBI 2014a).

The following section discusses how knowledge mining of the general literature, and of the National Health and Nutrition Examination Survey (NHANES) database in particular, is used to consider the potential impacts of environmental chemical exposures on the public health risks for diabetes. Multiple chemical exposures are also addressed.

3.2.1.1 Environment-wide Association Studies[30]

An environment-wide association study (EWAS) approach was used by Patel et al. (2012b) to investigate possible factors contributing to diabetes risks. In an EWAS, epidemiological data are comprehensively and systematically interpreted to identify the most important environmental factors associated with disease in a manner analogous to a GWAS. A GWAS associates genetic factors with disease on a genome-wide scale and has proper adjustment for the multiplicity of comparisons. An important difference is that an EWAS does not have a complete list of candidate environmental factors. Patel et al. (2012b) integrated genomic and toxicological data to identify genes, genetic variants, and environmental factors associated with type 2 diabetes. The method involved three steps:

1. Genetic and environmental data were summarized from VARIMED (VARiants Informing MEDicine, a genetic association database) and NHANES (an environmental exposure and effects database). VARIMED contains data on 11,977 gene variants, 9752 genes, and 2053 individuals; NHANES includes 261 genotyped loci, 266 environmental factors measured in blood and urine, and clinical measures for the same individuals.

2. Several environmental factors then were identified that positively or negatively affected risks for type 2 diabetes, including some nutrients and several persistent organic pollutants. Eighteen human genetic variations (SNPs) and five serum-based environmental factors also were identified that interacted in association with type 2 diabetes.

3. An analysis of the interactions among genes/gene variants and environmental factors with respect to risk for diabetes was conducted by Patel et al. (2013; 2012b).

Patel et al. (2013) report that the strongest evidence their analysis identified was for an interaction between rs13266634, a nonsynonymous coding SNP in the *SLC30A8* gene, and three nutrient factors, trans- and cis-b-carotene (which their statistical analysis indicated was associated with lower risk of diabetes) and c-tocopherol (which increased the risk). The *SLC30A8* gene is thought to modulate insulin, and Patel et al. (2013) hypothesized that impaired insulin secretion driven by the rs13266634 SNP might increase type 2 diabetes risk if combined with high or low levels of these specific nutrients.

The EWAS knowledge-mining method can be applied broadly to any number of common diseases to identify interactions between genetic and environmental factors and the impact on risks of disease. Patel and Cullen (2012) discuss a more comprehensive representation of chemical exposures

[30]This section is adapted largely from Patel et al. (2013; 2012b) with the assistance of Dr. Patel.

(termed the "envirome") and its use in evaluating the interplay of genetics and the environment. The EWAS approach is relatively new but has the potential to identify sensitive populations in response to exposures and to identify hypotheses or prioritize chemicals for their exposure-genome interactions. EPA will continue to evaluate its development and utility for Tier 2 assessments.

3.2.1.2 Expert Opinion

A recent National Toxicology Program (NTP) expert workshop considered evidence of causal associations between chemical exposures and increased risk of diabetes or obesity (Thayer et al. 2012). The data considered at the NTP workshop included approximately 870 findings from more than 200 human studies and the most useful and relevant endpoints from experimental animal and *in vitro* assays (e.g., ToxCast and Tox21 programs). Environmental factors considered at the workshop included maternal smoking and nicotine, arsenic, persistent organic pollutants, organotins, phthalates, bisphenol A, and pesticides. Overall, the results suggested that associations can be made between environmental factors and type 2 diabetes or obesity, although causality is more difficult to assign (Table 7). Mechanistic and *in vitro* studies played a demonstrable role in the evaluation of causality, particularly in the absence of traditional data.

3.2.1.3 Itemset Associations between Prediabetes/Diabetes and Chemical Exposures

As a complement to efforts by Thayer et al. (2012) and Patel et al. (2013; 2012b), we evaluated another data mining approach called frequent itemset mining using NHANES data following the approach described in Bell and Edwards (2014). Frequent itemset mining often is used with large, sparse data sets like financial transactions, shopping transactions, and, more recently, health care data. It can identify associations between a specific set of medical interventions and readmittance rates, or identify an item in which a shopper might be interested, based on his or her current cart. Bell and Edwards (2014) demonstrated the ability to use frequent itemset mining to identify and prioritize associations between environmental exposures and health effect markers using the NHANES data.

Building from earlier work (Bell, S. and Edwards 2014), we focused our analyses on the 2003–2004 and 2009–2010 NHANES cycles to identify associations between markers for diabetes and individual metals, examining single metal exposures in the 2003–2004 cycle and complex metal co-exposures in the 2009–2010 cycle. The Apriori algorithm (Agrawal et al. 1993; Borgelt and Kruse 2002; Hahsler et al. 2005) was used to identify association rules, X⟶Y, which identify items "Y" that are likely to co-occur with item "X." These rules carry with them parameters that help in interpreting their relevance (e.g., "support" and "confidence"). Support is the percentage of transactions (individuals or samples) in which all items in the rule are found. Confidence describes the proportion of people having "Y" that also had "X." So for example, in the rule diabetes⟶lead, the support would be the number of people who had both diabetes and elevated lead levels. The confidence in this case would be the number of people who had elevated blood lead out of all the people having markers for diabetes (percentage of the diabetic sample with elevated lead). A third parameter of interest described in Tables 8 and 9 below is "lift." Lift is the deviation of support for the rule from the expected support if both sides were independent. A lift of 1 implies that the two sides of the rule behave like random variables, a value less than 1 implies that the co-occurrence is

Table 7. Expert Judgment Concerning Causality for Diabetes/Obesity and Environmental Factors

Chemical/ Environmental Factor	Outcome	Conclusions from Breakout Group
Maternal smoking and nicotine	Childhood obesity	Likely causal, supported by epidemiology data and animal studies (Behl et al. 2013).
Arsenic	Diabetes	Sufficient evidence for a positive association between arsenic and diabetes in populations with relatively high exposure levels (≥150 µg arsenic/L in drinking water) (Maull et al. 2012).
Persistent organic pollutants	Diabetes	Sufficient evidence for a positive association of some organochlorine pollutants with type 2 diabetes (Taylor et al. 2013).
Organotins	Obesity	No human data; limited number of high quality animal, *in vitro*, and mechanistic studies of tributyl tin indicative of adipocyte differentiation (*in vitro* and *in vivo*); increased amount of fat tissue in adult animals exposed during fetal life (*in vivo*), and increased lipid accumulation in adipocytes and increased differentiation of multipotent stromal stem cells into adipocytes (*in vitro*) (Thayer et al. 2012)
Bisphenol A	Diabetes	Human data insufficient; primarily based on animal and *in vitro* studies, evidence is suggestive of an effect of bisphenol A on glucose homeostasis, insulin release, cellular signaling in pancreatic β cells, and adipogenesis (Thayer et al. 2012)
Phthalates	Diabetes or obesity	Human data insufficient; animal and human data suggestive of PPARα agonist activities of phthalate metabolites, species differences exist (Thayer et al. 2012)

less than expected, and a value greater than 1 implies a positive association. Our first study identified and ranked metals that were associated with the presence of prediabetes/diabetes markers (prediabetes/diabetes⟶chemical Y). Our second study examined the converse rules where chemical X⟶ prediabetes/diabetes. Only the top associations are presented.

Prediabetes/Diabetes and Individual Chemical Exposures

Table 8 lists the results of associations between prediabetes/diabetes and metal concentrations in blood or urine. These results suggest type 2 prediabetes/diabetes likely is associated with lead and cadmium (blood or urine) and possibly associated with arsenicals (urine). Type 2 prediabetes/diabetes is not likely associated with cesium and uranium alone. Taking the first row in the table as an example, 11 percent of the individuals from the NHANES 2003–2004 cycle who had shown markers for elevated metal exposure or markers for diabetes had both high blood lead and the presence of prediabetes/diabetes. Of individuals with markers for prediabetes/diabetes, 34 percent or roughly one-third had high blood lead levels. The strong lift value (1.44) implies a positive relationship and that they are likely not behaving independently.

Prediabetes/Diabetes and Multiple Chemical Exposures

Table 8. Top Metals Co-occurring with Type 2 Prediabetes/Diabetes Markers in NHANES 2003–2004

Metal Marker	Lift	Support	Confidence	Conclusion
High blood lead	1.44	0.11	0.34	Association
High urine cadmium	1.43	0.13	0.43	Association
High blood cadmium	1.26	0.09	0.30	Association
High urine arsenobetaine	1.25	0.10	0.33	Association
High urine lead	1.20	0.09	0.28	Association
High urine total arsenic	1.18	0.09	0.31	Association
High blood total mercury	1.12	0.09	0.30	Association
High urine cesium	1.03	0.08	0.25	No association
High urine uranium	1.01	0.07	0.24	No association

Using the 2009–2010 NHANES cycles, the association of prediabetes/diabetes markers, given co-occurrence of multiple metals in the body, was investigated (Table 9). These results support the findings in the NHANES 2003–2004 cycle analysis, with the strongest single associations in Table 8 exhibiting the strongest combined association in Table 9. Again taking the first entry in the table, support of 0.11 translates to 11 percent of the individuals with markers for diabetes or metals had elevated urine cadmium, blood lead, and urine total arsenic along with markers for diabetes. In this case, 59 percent of people with high levels of cadmium, lead, and arsenic also had markers for diabetes. Considering the large lift (1.46), an individual with elevated levels of lead, cadmium, and arsenic likely would be at risk for diabetes. The results in the NHANES 2003–2004 cycle analysis also point to some complex relationships whereby cesium, which was not strongly associated with the health effect markers in the Table 8 results, is considered associated if found in combination with other metals. Further work is needed to provide guidance in interpreting multiple-item associations with this type of analysis.

Synthesis of Frequent Itemset Mining Results

Overall, the frequent itemset mining results indicate that lead and cadmium exposure are highly likely to be associated with type 2 prediabetes/diabetes. High lead levels occurred in 9 of 10 and cadmium in 8 of 10 of the top-ranked rules in the multiple-chemical analysis of the data shown in Table 9. Further evidence is provided by the results where blood lead, blood cadmium, and urine cadmium were the highest rated outcomes based on lift in the single chemical analysis shown in Table 8. Confirmatory evidence exists that these metals might also be elevated in other diabetic populations (Afridi et al. 2008). Low dose mixtures of lead, cadmium, and arsenic might induce oxidative stress (Fowler et al. 2004), and evidence suggests that cadmium might induce hyperglycemia in rats (Bell, R. et al. 1990). The results of the two analyses (Tables 8 and 9) indicate that uranium and cesium are not likely to be associated with type 2 prediabetes/diabetes. Whether mercury is likely to be associated with type 2 prediabetes/diabetes remains unclear.

Based on this analysis, a large proportion (>50 percent) of the U.S. population with elevated lead, cadmium, and arsenic levels would be expected to have type 2 prediabetes/diabetes. These data are

Table 9. Strength of Association between Metal Co-exposures and the Presence of Diabetes/Prediabetes Markers in NHANES 2009–2010

Metal Markers	Lift	Support	Confidence	Conclusion
High urine cadmium High blood lead High total urine arsenic	1.46	0.11	0.59	Association
High urine cadmium High urine lead High blood lead High total urine arsenic	1.44	0.10	0.58	Association
High urine cadmium Low urine cobalt	1.40	0.11	0.56	Association
High urine cadmium High blood lead	1.38	0.17	0.56	Association
High urine cadmium High urine lead High blood lead	1.38	0.15	0.56	Association
High urine cadmium High urine cesium High blood lead	1.38	0.11	0.56	Association
High urine cadmium High blood cadmium High blood lead	1.37	0.13	0.55	Association
High urine lead High blood lead High total urine arsenic	1.37	0.12	0.55	Association
High urine cesium High blood lead High total urine arsenic	1.37	0.10	0.55	Association
High urine cadmium High urine lead High blood cadmium High blood lead	1.37	0.11	0.55	Association

not sufficient support for the hypothesis that these metals "cause" type 2 prediabetes/diabetes. They support only that these metals (or mixtures of these metals) are "associated" with type 2 prediabetes/diabetes.

This association could result from any the following: (1) the mixture of these chemicals do, in fact, cause type 2 prediabetes/diabetes; (2) prediabetic/diabetic phenotypes alter the absorption, distribution, metabolism, and excretion of these metals, and cause higher body burdens; (3) only one of these chemicals causes type 2 prediabetes/diabetes and leads to alterations in the absorption, distribution, metabolism, and excretion properties of the other chemicals; or (4) a

correlation with some other co-associated factor/exposure. Some data indicate that the three metals work together to induce oxidative stress. Other data suggest that cadmium itself might induce hyperglycemia in rats. Clearly, further studies are needed to resolve causality.

This exercise demonstrates the utility of frequent itemset mining to identify associations between chemical body burdens and potential disease endpoints and to provide direction for future studies needed to resolve the likely causal mechanisms for those associations. The results also illustrate how data mining methods from other disciplines can be applied to risk assessment, and provide valuable insight into associations between exposure and health effects.

3.2.1.4 Example: Characterizing Human Susceptibility and Population Variability

Risk managers can use genotype and allele frequency data in a nucleotide variations database (called dbSNP[31]) to understand population variance, and identify susceptible populations based on the underlying genetics, and chemical and nonchemical factors.[32] As an example, analysis of a random sample of 100 individuals of Mexican descent in Los Angeles found that 66 percent were homozygous for the risk allele for diabetes, 30 percent were heterozygous, and 4 percent were homozygous for the nonrisk allele (NCBI 2012). Assuming the sampling is representative of the entire population of Mexican-descended residents of Los Angeles, approximately 66 percent of these individuals might be at an increased risk of developing diabetes, independent of their body mass index (OMIM 2014). Heterozygous individuals (30 percent of the population) also might carry some risk and be affected by their zinc intake. Likewise, the heterozygous individuals might be more sensitive to other metals, chemicals, or dietary factors that could compete with zinc for absorption, or they might be more sensitive to chemicals that could interfere with zinc metabolism, transport, and insulin biosynthesis. Given the high rate of zinc deficiency in Mexican children that is not correlated with socioeconomic status, finding zinc deficiency in children of Mexican descent

---

[31]The **dbSNP** is world's largest database for nucleotide variations and is part of the National Center for Biotechnology Information (NCBI), an internationally respected resource for molecular biology information. As of this date, dbSNP comprises a large cluster of species-specific databases that contain more than 12 million nonredundant sequence variations (single nucleotide polymorphisms, insertion/deletions, and short tandem repeats) and more than 1 billion individual genotypes from HapMap and other large-scale genotyping activities—more than 200 GB of data and growing daily.

[32]Gene-disease associations can be identified using a combination of EWAS and GWAS. The work by Patel et al. (2013) demonstrates the use of an EWAS to identify potential interactions among SNPs (i.e., a mutation of a single nucleotide within the DNA of a gene sequence), environmental chemical levels in blood and urine, and health effects—specifically type 2 diabetes—using data from NHANES. Although support for genotype and chemical interactions was limited, interesting interactions were noted between the nonsynonymous coding SNP rs13266634 in the *SLC30A8* gene and cis- and trans-beta-carotene and gamma-tocopherol. This SNP has been associated with type 2 diabetes previously (Diabetes Genetics Initiative of Broad Institute of Harvard et al. 2007; Pare et al. 2008; Rung et al. 2009; Scott et al. 2007; Sladek et al. 2007; Steinthorsdottir et al. 2007; Takeuchi et al. 2009; Timpson et al. 2009; Zeggini et al. 2007). The *SLC30A8* gene is a zinc transporter found in the pancreatic beta-cell secretory vesicles. Zinc has been associated with insulin biosynthesis (Emdin et al. 1980), and chronic decreased zinc intake has been associated with an increased risk of diabetes (Miao et al. 2013). The risk allele in rs13266634 is C (Sladek et al. 2007), while the minor allele is T (NCBI 2012). (Note that single genes and variants of that gene, and the relationship to disease, are often studied in isolation, when many genes, in fact, might contribute to the risk of disease.

living in Los Angeles might not be surprising, especially if diet plays a significant role in the deficiency (Morales-Ruan et al. 2012).

We also can hypothesize that cadmium exposure will be of concern to individuals who are homozygous or heterozygous for the risk allele. Cadmium has been shown to compete with zinc transporters and might lead to beta-cell dysfunction, lack of insulin production, and ultimately hyperglycemia (El Muayed et al. 2012). Individuals with the rs13266634 risk allele could be more sensitive to cadmium exposures than the rest of the population.

Through database mining and an understanding of the pathway affected by the allele and a chemical's AOP network, we can identify potentially susceptible populations more easily. This example could be extended by examining cadmium exposure data for the Los Angeles area and using a geographic information systems approach with census data to identify potentially susceptible individuals, based on the allele probabilities. This type of predictive modeling could help advance risk management with more definitive and targeted community-level responses.

3.2.2 Short-term *In Vivo* Bioassays – Alternative Species

Short-term *in vivo* bioassays using alternative species (i.e., nonmammalian species) provide data to identify hazards, integrate dose-response effects, and understand pathways and resulting adverse health effects. The types of alternative or nonmammalian species (e.g., zebrafish, yeast) used in scientific exploration can vary widely. Considerable toxicological work has been completed on fish, but work in very simple organisms such as yeast also provides insight into cellular regulation at multiple levels that control core biological processes and enable cells to respond to genetic and environmental changes (Yeung et al. 2011). These assays are useful for assessing chemical risks to humans and other species.

Four advantages of using *in vivo* assays with alternative species in contrast to using *in vitro* assays include:

1. full representation of the normal metabolic capability of the species under study;
2. evaluation of complex phenomena, such as birth defects or neurobehavioral alterations, effects requiring fully functional tissues, and cell-to-cell or tissue interactions;
3. as a function of 1 and 2 above, molecular changes and phenotypic outcomes can be studied rapidly and relatively inexpensively in the same organism; and
4. alternative species *in vivo* assays are faster and relatively inexpensive to perform over the full lifespan of the organism (relative to mammalian species), facilitating study of the entire disease etiology, from the MIEs to adverse health effects, due to the shorter lifespans.

Studies in nonmammalian species are playing a progressively more important role in chemical testing, hazard identification, and dose-response assessment for both nonhumans and humans (EC 2011; ECHA 2013; EPA 2012c; OECD 2004b; Perkins et al. 2013; Schug et al. 2011; Vacaru et al. 2014). Both the European Chemicals Agency (ECHA) and EPA consider nonmammalian species tests in the study of endocrine disruptors (ECHA 2013; EPA 2009b, 2014h) to evaluate, in this case,

environmental risks rather than human health. EPA has indicated its intention also to include alternative species in the Endocrine Disruptors Screening Program to evaluate risks to human health (EPA 2011c). As a caveat, although nearly all important biological functions are preserved across species, the exact relationships between molecular functions and phenotype outcomes have not always been preserved. Additionally, calculation of exposure-dose across species and routes of exposure present challenges. Thus, species-to-species extrapolation remains an important risk assessment challenge and area of active research.

The following prototype demonstrates how alternative species studies might be used for prioritization and screening or as the basis for Tier 2 type assessments. Specifically, the prototype example examines use of alternative species to identify and characterize thyroid hormone disruption.

3.2.2.1 Tier 2 Prototype: Using Alternative Species to Identify Thyroid Hormone Disruption

Endocrine disrupting chemicals (EDCs) are chemicals that interfere with endocrine hormone signaling and produce adverse effects in both humans and wildlife.[33] In a state-of-the-science review, the World Health Organization (WHO) concluded that thyroid disruption-associated neurobehavioral disorders are occurring in children, and the incidence of disorders has increased in recent decades (WHO 2012). Normal thyroid function is essential for normal brain development, during fetal and early childhood development. Thyroid hormones are also crucial to inner ear and bone development, and to bone remodeling and physiological functions such as metabolism (De Coster and van Larebeke 2012). Internationally agreed-upon and validated test methods for identification of endocrine disruptors address only a limited range of the known endocrine disrupting effects (Miller et al. 2009). In its state-of-the-science review, WHO advised that existing testing protocols do not characterize all essential functions completely and that adverse effects "are being overlooked" (WHO 2012).

In testing for potential EDCs, the role of the thyroid hormone is of particular toxicological interest because the dependence of post-embryonic development on thyroid hormones is a common feature of vertebrate ontogeny (Paris and Laudet 2008). Human and vertebrate post-embryonic neurodevelopment is thyroid hormone dependent and deviations from normal thyroid hormone

[33]Endocrine hormones are secreted internally from glands, and distributed in the body via the bloodstream. The best-known endocrine hormones are the sex hormones, estrogens and androgens, and the thyroid hormones. Hormones act as signals to help orchestrate several development, reproductive, and growth functions. They are released in response to various internal and external stimuli, and travel throughout the body at very low levels (parts per billions) until they bind to receptors on cell surfaces and stimulate their intended intracellular response. Disruption of hormone signaling can occur from external exposures to EDCs that act as hormone receptor agonists or antagonists, or that interfere with hormone production or kinetics (release, transport, metabolism, excretion). These disruptions can produce profound adverse effects in the many biological processes controlled or influenced by endocrine hormones. Specific effects associated with EDCs include learning disabilities, severe attention deficit disorder, cognitive and brain development problems; deformations of the body (including limbs); breast cancer, prostate cancer, thyroid and other cancers; and sexual developmental dysregulation such as feminizing of males or masculinization of females.

concentrations at critical times are associated with a variety of neurological defects and deficits (Zoeller et al. 2002). This period of development typically is characterized by transient elevations of thyroid hormone that elicit species-specific physiological and morphogenetic responses with lasting developmental consequences. Although outcomes might differ among species, thyroid hormone regulation is generally essential for normal development in vertebrates, thereby establishing the basis for cross-species extrapolation of developmental risks. Several methods using alternative species have been proposed to measure these outcomes for thyroid pathways (Makris et al. 2011; Nichols et al. 2011). The timing (or window) of exposure is critical, as the impact of thyroid hormones changes as the brain develops (Zoeller and Rovet 2004). Transitions from tadpoles to juvenile frogs and body plan reorganization in flatfish are two nonmammalian examples of thyroid hormone-controlled events.

A key factor in thyroid hormone-related risk assessment is the ability to examine hormone disruption and the resultant developmental disruption at higher levels of tissue organization. Results from omics technologies and other thyroid hormone toxicity assessments, such as EPA's ToxCast chemical screening efforts (EPA 2008), can be linked to adverse outcome data from alternative species studies. Two examples are:

1. construction of regulatory networks using time-series data in genotyped populations and integration of multiple data types (e.g., endogenous metabolite concentrations, RNA expression, DNA variation, DNA-protein binding); and

2. chemicals identified as potential developmental disruptors in high-throughput screening (HTS) assays that are further evaluated with available *in vivo* effects data to establish dose-response relationships, windows of susceptibility, potential impacts of maternal exposure on progeny, and existence of subtle impacts on behavior, learning, and memory.

3.2.2.2 System and Pathway Models

As discussed throughout this document, a systems biology perspective in understanding the events leading to an adverse effect is central to the use of molecular biology data in risk assessment. In the absence of an organizing mechanistic concept or anchoring to traditional data, interpretation of omic changes is highly uncertain and in general unsuitable for risk assessment other than prioritization and screening for additional work. Pathway analyses are useful to inform extrapolation across species and to aid in characterizing the variability within populations through identifying and describing both initiating, and other, key biological events leading to adverse outcomes. They also can help identify how human-focused screening data can inform ecological risk assessment. Although making quantitative predictions of disease risks based on today's system biology or adverse outcome models is often very difficult, progress is being made, and pathway analysis remains a top priority for advancing dose-response assessment.

Thyroid hormone disruption can occur at many points in a complex process and at different levels of biological organization. Figure 23 illustrates different ways that different classes of chemicals can disrupt thyroid hormone regulation and signaling. In humans, disruption leads to birth defects, decreased IQ, and metabolic disorders, and potentially to cancer. In rats, increased thyroid-stimulating hormone (TSH) leads to thyroid hyperplasia and cancer. Understanding the system as a

whole provides the most useful risk information, including increased evidence for hazard identification and dose-response assessment, characterization of population variability, cross-species extrapolation, and evaluation of mixtures.

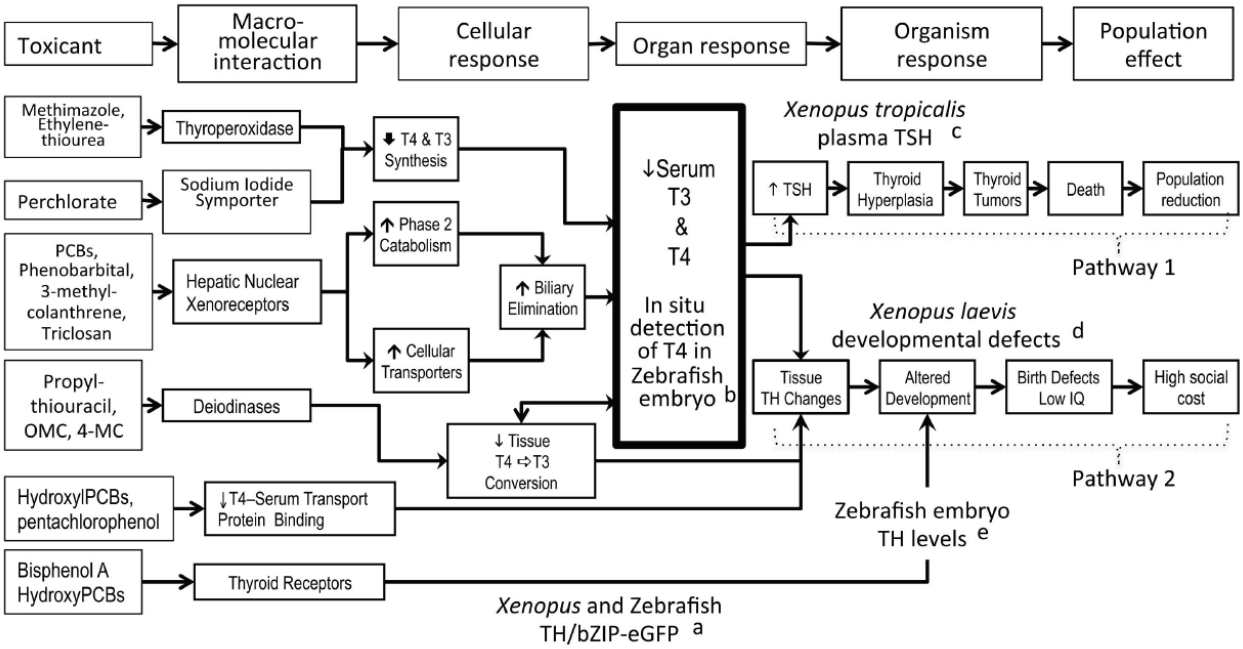


Figure 23. Major Adverse Outcome Pathways (AOPs) for Thyroid Disruption with Example Toxicants and Alternative Models Applicable to Both Human and Ecological Hazard Assessment (Perkins et al. 2013).
Reproduced with permission from *Environmental Health Perspectives*.
The thick black outlined box indicates the critical event of serum level concentrations of thyroid hormones. Pathway 1: rat pathway leading to tumors via thyroid hyperplasia. Pathway 2: principle pathway of concern affecting humans. Abbreviations: IQ, intelligence quotient; 4-MC, 4-methylbenzylidene camphor; OMC, octyl methoxycinnamate; T_3, triiodothyronine; T_4, thyroxine; TR, thyroid receptor. Figure modified from Crofton (2008) .[a]Quantification of plasma TSH levels in *Xenopus tropicalis* (Korte et al. 2011). [b]Direct quantification of intrafollicular concentrations of T_4 in zebrafish embryos (Thienpont et al. 2011). [c]Detection of developmental defects with *X. laevis* metamorphosis assay (Degitz et al. 2005; OECD 2004a). [d]Detection of developmental defects using zebrafish embryos. [e]Reporter gene (eGFP) detection of TR activity (Fini et al. 2007).

3.2.2.3 Dose-Response Relationships for Human Disease

Although quantitatively predicting human disease risks is currently difficult, several approaches using alternative species provide information on causal mechanisms as well as data on the potency of chemicals that cause effects. Examples of these approaches include the use of biomarkers of exposure and effect, assessments of relative potency to induce adverse effects, species extrapolation, and benchmark analysis to characterize the dose-response relationship.

Biomarkers of Exposure and Effects

Key events in the perturbed pathway can be represented with biomarkers of exposure and effect. In situations where considerable systems biology information links the event to outcomes, a biomarker might provide a measure of hazard for risk assessment. For example, upstream events in thyroid hormone pathways converge on serum levels of the thyroid hormones, triiodothyronine (T3) and thyroxine (T4). Downstream events occur in peripheral tissues where a significant degree

of species-specific effects is observed (Figure 24). Thus, serum T4 level can be used as a biomarker of thyroid function across species. In the laboratory, researchers use T4 and TSH levels in fish and frogs to assess the thyroid disrupting potential of chemicals (Thienpont et al. 2011; Tietge et al. 2013). To assess human exposures, the Centers for Disease Control and Prevention (CDC) has used decreased serum levels of T4 and increased levels of TSH measured in the U.S. population to infer increased potential risks for thyroid dysfunction-related disorders at low levels of perchlorate exposures (Blount et al. 2007; Lau et al. 2013).

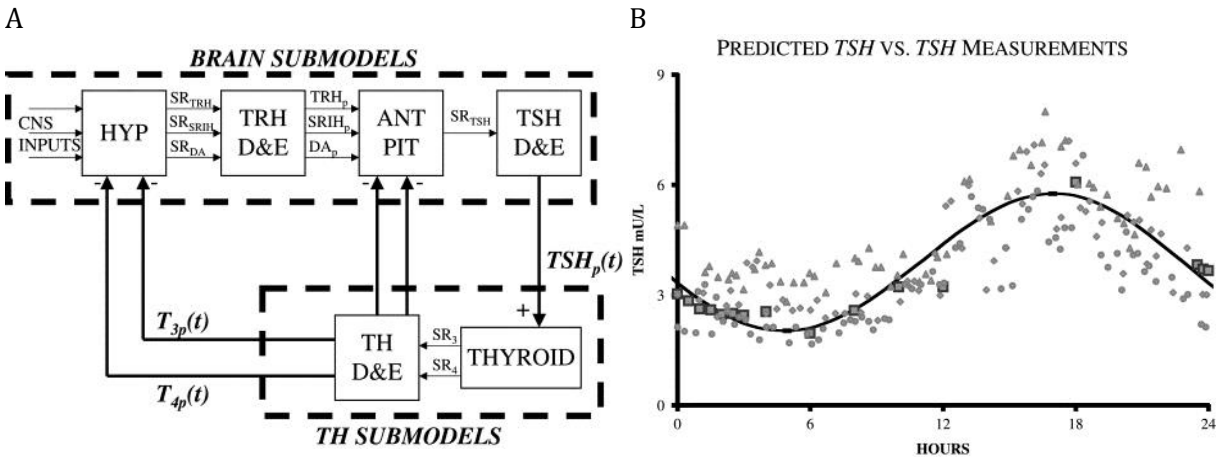


Figure 24. Dose-response Relationships.
Within species, significant advances are being made in quantitative systems biology modeling (Eisenberg et al. 2008). Panel A: Overall feedback control system model of thyroid hormone regulation with three source organ blocks (hypothalamus [HYP], anterior pituitary [ANT PIT], and thyroid glands [THYROID]); three sink blocks (for TRH, TSH, and T3 and T4 distribution); and elimination (elimination = metabolism and excretion) (D&E). TRH = thyrotropin-releasing hormone; TSH = thyroid-stimulating hormone; T3 = triiodothyronine; T4 = thyroxine; SR = secretion rate; p = plasma or portal plasma for TRH-related components; DA = dopamine; SRIH = somatastatin. Panel B: Feedback control system (FBCS) model validation study results. Predicted normal circadian TSH versus independent TSH data (not used in fitting the FBCS model) (triangles and diamonds represent data from Sarapura et al. (2002), circles represent data from Samuels et al. (1994). Also shown (squares) are the mean TSH data from the larger database used to fit the FBCS model of Blakesley et al. (2004). Reproduced with permission from *Mary Ann Liebert, Inc*.

Relative Potency

Identification of pathways and assays impacted by chemicals can be useful in initial prioritization of many compounds. Potentially toxic chemicals can be identified through predictive models built on relationships between *in vitro* ToxCast assay results and *in vivo* effects, as demonstrated in an analysis identifying developmental toxicants (Sipes et al. 2011b). Focused *in vivo* tests with alternative species provide additional dose-response data and information about exposure window-response relationships. Pathway effects defined through gene expression changes can be

used to define a benchmark dose (BMD)[34] to characterize the sensitivity of an animal to a chemical (Thomas, R. S. et al. 2011).

Alternative species *in vivo* test systems can detect effects from mechanisms not represented in the *in vitro* high-throughput screening (HTS) assays. As an example, zebrafish were used to assess the 309 EPA ToxCast Phase I chemicals for potential developmental toxicity to both humans and ecological species. In fish embryo or larval cultures, 191 (62 percent) chemicals were toxic (death or malformations) to the developing zebrafish. Both toxicity incidence and potency were correlated with chemical class and hydrophobicity. As an integrated model of the developing vertebrate, the zebrafish embryo screen provides information relative to overt and organismal toxicity. In 12 classes of chemicals, 100 percent of the chemicals induced developmental toxicity, 4 classes of which induced developmental toxicity with an AC_{50}[35] below 4 µM. Translating such results directly into dose-response for human risks is difficult, but Padilla et al. (2012) argue that alternative species can be used to build relative rankings of chemicals based on their potency to cause adverse effect. Such rankings can be used to prioritize chemicals or classes of chemicals for additional evaluation.

Characterizing the Dose-Response Relationship

Chemical dose-response relationships characterized in one species can be extrapolated to other species including humans, if sufficient pathway-based data and kinetic information are available (Perkins et al. 2013). Because many biological functions and pathways are conserved across species, similarity of genes encoding those pathways provides support for direct comparisons of pathway or genomic effects among species. Where pathways are highly conserved, the dose-response relationship in the alternative species can be extrapolated to an analogous pathway in mammals. For example, pathways in the hypothalamus-pituitary-gonad (HPG) axis are highly conserved among vertebrates. Based on similarities in the HPG pathways, the chemical effects in fathead minnows have been shown to be predictive of endocrine disrupting effects in rats (Ankley and Gray 2013). Qualitative predictions of hazard are likely to be tenable based on similarities in hypothalamus-pituitary-thyroid (HPT)-dependent pathways among species (i.e., iodine uptake), however, even though altered iodine uptake hinders development, the most sensitive outcome indicator might be different among species. In rats, for example, thyroid hormone disruption can lead to thyroid tumor development (Hurley 1998), while in frogs, metamorphosis is disrupted (Degitz et al. 2005). Dellarco et al. (2006) further discuss some of the challenges to cross-species extrapolation.

---

[34]**Benchmark dose (BMD)** is a quantitative value that describes the dose-response relationship based on a model that incorporates all of the dose-response data. BMD is the dose that is expected to result in a specified percent (called the benchmark response or BMR level) of the population exhibiting the adverse effect(s) associated with chemical. BMR is generally set near the low end of the observable range of the data, generally at an incidence rate of around 5–10% incidence (EPA 2012a).

[35]**AC_{50}** is the concentration at which activity is 50% of its maximum. This value is useful in comparing assay results.

Although species differ in absorption, distribution, metabolism, and excretion of chemicals, species-to-species dose extrapolation is possible. Differences across species and routes of exposure are important considerations when extrapolating data from alternative species to humans. Considerable experience has been gained in developing physiologically based pharmacokinetic (PBPK)[36] models for extrapolating dose among mammalian species (Mumtaz et al. 2012; Thompson et al. 2008). Based on this experience, useful kinetic models also can be developed to conduct dose extrapolation from nonmammalian species. For example, concentrations in fish plasma from aqueous exposures have been extrapolated to a dose that yields an equivalent concentration in human plasma using an appropriate kinetic model (Schreiber, R. et al. 2011). As more information becomes available on the kinetics of chemicals in *in vivo* assays using newer alternative species, kinetic models will be developed to support the species-to-species extrapolations of dose for a variety of dosing regimens.

Normal post-embryonic development that depends on thyroid hormones requires coordinated spatial-temporal control of thyroid hormone activity. Such activity is regulated not only through the classical features of the HPT axis, but also through peripheral mechanisms external to the hypothalamus, pituitary, and thyroid tissues, such as differential regulation of deiodinase activity, hepatic metabolism and excretion of thyroid hormones, thyroid hormone receptor regulation, and transmembrane thyroid hormone transport. Of these major controlling processes, the mechanisms of the central HPT axis are generally considered conserved across vertebrate species, and useful for comparative efforts; those of the peripheral tissues, however, are typically more divergent and must be used with care in cross-species analysis.

3.2.2.4 Population Variability

Understanding the variation of an individual relative to population variation can be key to identifying an adverse effect on a population. Polymorphisms affecting drug responses can vary widely in populations. In humans, 20–25 percent of prescription drugs are metabolized in the liver by cytochrome P450 CYP2D6, where variants confer widely different rates of drug metabolism, such that some people might respond with an onset of toxicity while others fail to experience efficacy (Ingelman-Sundberg 2005). Variants causing unanticipated results can comprise a significant portion of a population, and that distribution can vary widely across populations (Andersen, S. et al. 2002; Ingelman-Sundberg 2005; Sistonen et al. 2007; Wooding et al. 2002). Understanding the variation in adverse responses across a diverse testing population helps reduce the uncertainty of extrapolating laboratory data to real populations. In ecological risk assessment, differential response to chemicals is an important consideration where not only potentially sensitive subpopulations might exist, but also sensitive species.

---

[36]**Physiologically based pharmacokinetic (PBPK)** models simulate pharmacokinetics in the body and are used to estimate the dose to a target tissue or organ by accounting for the rates of absorption, distribution among target organs and tissues, metabolism, and excretion. PBPK models also are often referred to as physiologically based toxicokinetic (PBTK) models in risk assessment to clearly distinguish the chemical as a toxicant (IRIS glossary; EPA 2014o). Both terms are in common use, and might appear in the text of this document. They relate to the same kind of model and are interchangeable.

Approaches have been developed to incorporate population diversity into toxicity testing through the use of large collections of different genetic lines of mice or cell cultures derived from them (Harrill et al. 2009; O'Shea et al. 2011; Rusyn et al. 2010). Alternative species could be especially useful for incorporating population variability into toxicity testing. The diversity in laboratory lines and outbred populations of fish can be high, especially if populations are collected from different areas impacted by pollutants (Guryev et al. 2006; Williams and Oleksiak 2011). Divergent lines of zebrafish can be used to examine variation in responses to chemicals in addition to determining possible genetic factors influencing adverse effects. As an example, Waits and Nebert (2011) crossed zebrafish lines displaying different levels of sensitivity to dioxin-like chemically induced developmental cardiotoxicity. The crosses were used in genome-wide quantitative trait loci mapping to identify several genes that contribute to the gene-gene and gene-environment interactions (in addition to the AhR). Their results demonstrated that chemically induced cardiac teratogenicity was a multifactorial complex trait influenced by gene-gene and gene-environment interactions, and that the identified quantitative trait loci harbor many dioxin-like chemically responsive genes critical to cardiovascular development. This approach provided useful new insights into the genetic basis of susceptibility to AhR-mediated developmental toxicity.

Although genetic diversity can be incorporated into testing using a panel of genetically inbred lines, unexpected results can occur. In a study comparing the responses of 19 inbred to 20 outbred zebrafish lines, Brown et al. (2011) found that effects of the EDC clotrimazole were dramatically different. Clotrimazole acts by inhibiting P450 activities involved in steroidogenesis production in fish. In inbred fish lines, 11-ketotestosterone production via steroidogenesis was significantly inhibited. In contrast, outbred lines responded with Leydig cell proliferation in testes and normal plasma concentrations of 11-ketotestosterone indicating that the outbred lines could compensate for inhibition by clotrimazole. Here, inbreeding had a strong impact on the diversity and type of response to the endocrine disruptor.

Overall, several new approaches are available that can help with better characterizations of population variability. These include the use of (1) AOP networks for identifying chemicals and other environmental stressors that appear to act by the same mechanisms and could contribute to risk; (2) *in vivo* and *in vitro* test results from genetically diverse populations for capturing the range of genetically determined risk; and (3) epidemiology studies for capturing variability due to molecular biological differences in response to chemical and nonchemical stressor exposures.

3.2.2.5 Cumulative Risks

As has been described elsewhere in this document, correct identification of causal perturbations that lead to adverse outcomes will enable determination of which environmental factors are likely to contribute to the cumulative risk for specific outcomes and which are not. Additionally, testing of combinations of chemicals can be conducted efficiently in alternative species. For example, alterations in neurosensory functions and intrafollicular thyroxine content of zebrafish exposed to potential disruptors have proven to be useful tools for evaluating multiple chemicals (Froehlicher et al. 2009; Raldua et al. 2012; Thienpont et al. 2011), as has the zebrafish developmental assay. Figure 25 illustrates the toxicity of chemical classes in the zebrafish developmental assay data (Padilla et al. 2012). Also available, but not shown here, are the dose-response data for each of the

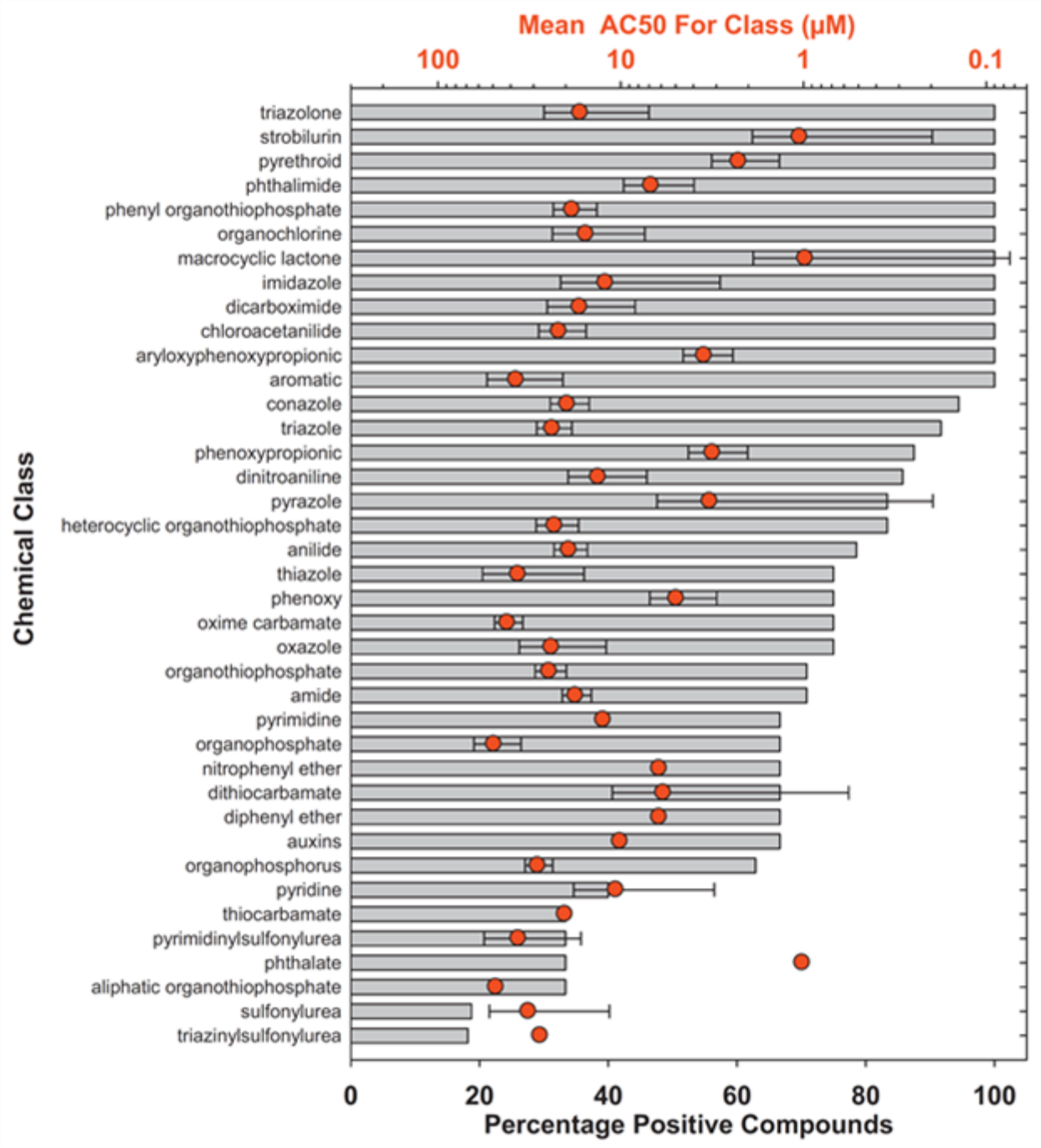


Figure 25. Relationship between Chemical Class and Toxicity to Developing Zebrafish for 300+ Chemicals.
The percentage positive chemicals in each class are represented by the gray bars (bottom axis), and the average AC_{50} for each group (±SEM) is indicated by the filled red circles (top axis). Only classes with three or more total members were analyzed, and only classes with at least two positive chemicals were included in the graph. If a class had only two positive chemicals, no error bars are shown (i.e., triazinylsulfonylurea, aliphatic organothiophosphate, phthalate, thiocarbamate, auxins, diphenyl ether, nitrophenyl ether, and pyrimidine) (Padilla et al. 2012). Reproduced with permission from *Elsevier*.

more than 300 chemicals that comprise the chemical classes. Thus, using these types of data, evaluation of both the individual chemical and the chemical class is enabled.

3.2.3 Short-term *In Vivo* Bioassays – Rodents

The use of short-term *in vivo* mammalian bioassays to support Tier 2 assessments is described here. The prototype example is based on research described in papers by R.S. Thomas et al. (2011) and discussed further in R.S. Thomas et al. (2013c; 2013d) on the use of short-term mammalian *in vivo* transcriptomic assays to predict chemical toxicity and dose-response (see Box 7 about the "transcriptome"). This research was a NexGen collaborative effort between EPA and The Hamner Institutes for Health Sciences. Female B6C3F1 mice were exposed to multiple concentrations of five chemicals found to be positive for lung or liver tumor formation in a 2-year rodent cancer bioassay (Thomas, R. S. et al. 2011; Thomas, R. S. et al. 2012c). Histological and organ weight changes were evaluated and gene expression microarray analysis was performed on the liver or lung tissues. The histological changes, organ weight changes, and tumor incidences in traditional bioassays were analyzed using standard BMD dose-response modeling methods to identify noncancer and cancer points-of-departure. The dose-related changes in gene expression were analyzed using a modification of EPA's BMD approach (EPA 1995). The analyses in R. S. Thomas et al. (2013c; 2013d) correlated the lowest transcriptional BMD with a cancer or noncancer BMD that had been identified from the traditional toxicity study data, rather than attempting to predict an apical effect based on an affected pathway. Efforts to explore the underlying mechanism were limited to grouping gene expression changes based on both biological processes and canonical signaling pathways. A comparison of the transcriptional BMD values with the traditional noncancer and cancer endpoint BMDs (see Figure 26) showed a high degree of correlation for specific biological processes (Thomas, R. S. et al. 2011) and signaling pathways (Thomas, R. S. et al. 2012c). In addition, transcriptional changes in the most sensitive pathway were also highly correlated with the adverse health effects observed in the traditional *in vivo* studies.

Box 7. What is the Transcriptome?

Ribonucleic acid (RNA) is the functional outcome of deoxyribonucleic acid (DNA) transcription, which is regulated by transcription factors. Researchers study the transcriptome the set of all RNA molecules in a given cell to identify gene expression patterns, or signatures. Specifically, short term transcriptomic assays in mammalian and alternative species enable observations of the effects of chemical exposure across multiple tissues.

The effects of exposure duration on outcomes are a key issue in the design and use of these new types of bioassays. Further studies demonstrated the consistency of the correlation between transcriptional changes and adverse health effects across different exposure periods (5 days to 13 weeks) (Thomas, R. S. et al. 2013d). The results shown in Figure 26 indicate that overall, the BMDs based on the transcriptional assay data are lower than those derived from the traditional assay data, and that the transcriptional BMDs could serve as a health-protective indicator of biological activity.

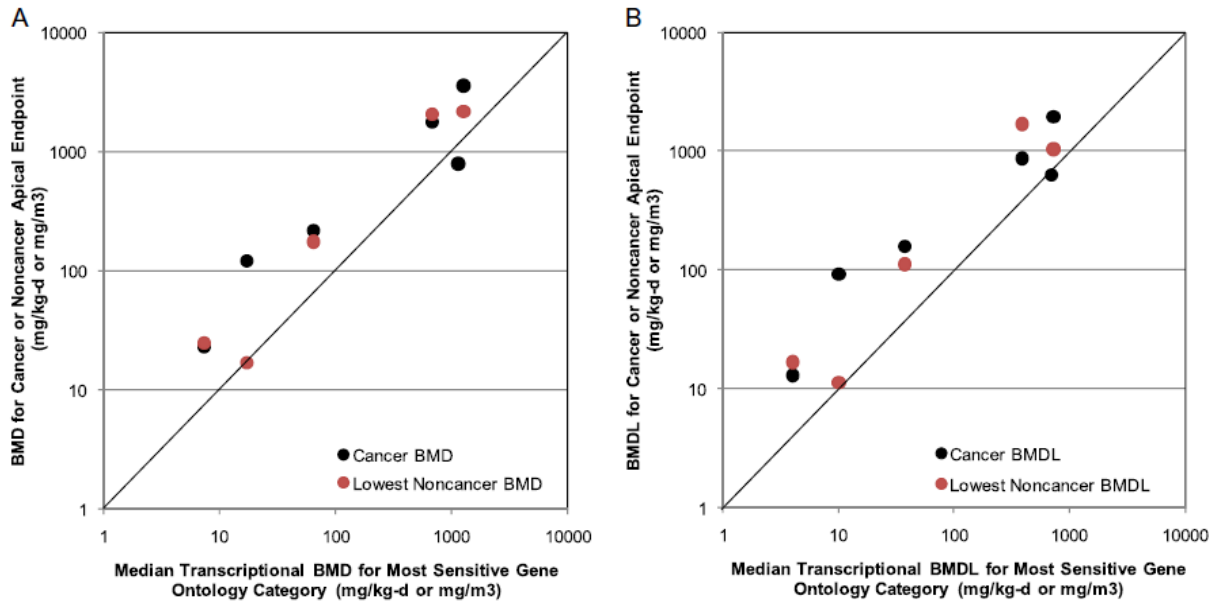


Figure 26. Scatter Plot of the Relationship Between (A) Benchmark Dose (BMD) and (B) Benchmark Dose Lower Limit (BMDL) Values for the Cancer and Noncancer Endpoints and the Transcriptional BMD and BMDL Values for the Most Sensitive Gene Ontology Category.

BMDL is a statistical lower confidence limit on the dose at BMD. For each chemical and tissue, BMD and BMDL values for tumor incidence and the lowest noncancer BMD and BMDL values were plotted. No noncancer BMD or BMDL values were plotted for the proteasome subunit, methylene chloride (MECL), in the lung because of the absence of histological changes (Thomas, R. S. et al. 2011). Reproduced with permission from *Oxford Journals*.

3.2.3.1 Hazard Identification

Short-term *in vivo* transcriptomic assays provide the metabolic capability and systems-level integration of whole-animal studies with a more rapid assessment of response to chemical treatment based on molecular-level data. A host of previous studies has demonstrated that gene expression signatures from short-term *in vivo* studies can be used to predict both subchronic and chronic toxic responses (Auerbach et al. 2010; Ellinger-Ziegelbauer et al. 2008; Fielden et al. 2011; Fielden et al. 2007; Fielden et al. 2005; Fielden et al. 2008; Nie et al. 2006; Thomas, R. S. et al. 2009; Thomas, R. S. et al. 2007; Thomas, R. S. et al. 2013d; Uehara et al. 2011). A transcriptomic "signature" typically is defined as a subset of genes for which the qualitative or quantitative expression pattern can be used to predict an *in vivo* adverse response with a defined accuracy. This approach remains relatively new, and more short-term *in vivo* transcriptomic data, standardized study designs, and identification of gene expression patterns and network perturbations are needed to advance our ability to predict chemical toxicity comparable to longer term assays. Dellarco et al. (2006) discuss some of the key challenges in correlating transcriptomic data with histopathology data, the traditional "gold standard" for characterizing adverse effects.

To develop a broad-based repertoire of gene expression signatures for hazard prediction, several factors are worth considering. First, the number of endpoints included should be sufficient to enable a comprehensive prediction of toxicological hazard. Previous studies that have used gene expression microarray analysis following short-term exposures of chemicals have been limited in the breadth of endpoints examined. These endpoints include the prediction of rat liver tumors

(Auerbach et al. 2010; Ellinger-Ziegelbauer et al. 2008; Fielden et al. 2011; Fielden et al. 2007; Fielden et al. 2008; Nie et al. 2006; Uehara et al. 2011), mouse lung tumors (Thomas, R. S. et al. 2009), and rat renal tubular toxicity (Fielden et al. 2005). A more comprehensive strategy would be to select a battery of tissues that includes those most frequently positive in rodent cancer bioassays (i.e., liver, lung, mammary gland, stomach, vascular system, kidney, hematopoietic system, and urinary bladder) and tissues commonly affected by noncancer disease. In a previous analysis, cancers in the eight tissues in parentheses above accounted for the observed cancers from exposure to 92 and 82 percent of all mouse and rat carcinogens, respectively (Gold et al. 2001). Additional tissues would be needed to be included for developmental and reproductive effects (including tissues from the developing fetus and reproductive organs).

Second, the number of positive and negative chemicals for each endpoint needs to be sufficiently large to support defensible conclusions, and the chemical diversity should represent the diversity in the domain of chemicals that need to be assessed for potential effects. For complex toxicological responses such as tumor formation, a previous study estimated that at least 25 chemicals were necessary (Thomas, R. S. et al. 2009).

Third, selection of the time point to perform the gene expression analysis is an important decision. The time point selection is a balance between cost (i.e., the shorter the time point, the less expensive the study) and a more stable gene expression signature. Among the previous efforts, certain studies relied on much shorter time points (e.g., 5 days), but tended to increase the dose beyond that which would be tolerated in a chronic bioassay (Fielden et al. 2007). Other studies used the same doses as those in the chronic bioassay, but used exposures longer than 5 days (Thomas, R. S. et al. 2009). In one study that examined the effect of exposure duration, the overall conclusion was that increasing exposure duration (2–90 days) increased the predictive performance of the gene expression signatures for genotoxicants (Auerbach et al. 2010).

3.2.3.2 Exposure/Dose-Response Assessment Using High-throughput Screening (HTS)

With the advent of HTS, the potential to screen thousands of chemicals for biological activity presents as many challenges as promises. If HTS can decrease the number of chemicals of interest by 90 percent (a 10 percent hit rate across chemicals and assays), the resulting number still would overwhelm the throughput of the traditional toxicity-testing paradigm. Clearly, a multi-tiered approach to prioritization can lead to more effective applications of animal toxicity testing. The development of predictive gene expression signatures and dose-response studies would provide a relatively efficient and cost-effective method for both identifying chemicals of concern and estimating a point of departure for adverse responses. This information would help support large-scale prioritization and regulatory efforts in the United States and Europe. The gene expression data combined with other data types (e.g., toxicity data from similar chemicals, pharmacokinetic [PK][37] data) could provide sufficient information to evaluate toxicity. Confidence in the evaluation,

[37]**Pharmacokinetics (PK)** – The root word ''pharmakon'' has complex meaning that encompasses both a remedy and a toxicant (and more broadly any biologically active substance). Risk assessors sometimes use

however, would depend on the overall strength of the evidence. Expression changes can vary depending on dose, time, species, tissue lifestage, and individual genetic profile. Such changes increase the complexity of identifying causal relationships between exposures, specific signatures, and outcomes.

3.2.4 Risk Assessment Implications Across the Tier 2 Prototypes

The three Tier 2 prototypes described above illustrate different approaches that are intermediate in resource use and scope between the more robust and resource-consuming approaches for Tier 3 major-scope assessments, and the high-throughput (HT) and cost-efficient approaches discussed in the next section for Tier 1 screening and prioritization assessments. The development of the Tier 2 prototypes led us to the following inferences.

3.2.4.1 Knowledge Mining

- New knowledge mining approaches, as performed for the diabetes and obesity prototype, rely on computer-assisted implementation of algorithms that identify, integrate, and interpret large amounts of data. These approaches take advantage of the huge, relatively new, databases managed by NIH and others into which almost all new published data are placed. Hence, a large swath of existing literature can be brought to bear on environmental problems in an unprecedented manner.
- The computerized component of knowledge mining is in essence high throughput (i.e., the literature and data for thousands of chemicals screened in an automated way over a short time). The human components of quality assurance and interpretation are the elements that add time and resources, making this a Tier 2 prototype method rather than an exclusively HT approach. Automated knowledge mining can provide more information that is quantifiable than traditional literature searches and can integrate information across diseases, chemicals, and other risk factors for further evaluation.
- Information acquired by knowledge mining is primarily associative in nature, hence, most suitable for hypothesis generation and in screening and prioritization. Use in identification of hazards and toxicity values would be suggestive at best. Additional meta-analyses of multiple epidemiological, experimental, and mechanistic studies can add to the weight of evidence for potential risks identified by knowledge mining, thus potentially expanding its use in risk assessment.

3.2.4.2 Short-duration *In Vivo* Exposure Paradigms

The other two Tier 2 prototypes evaluated use of short-duration, *in vivo* data from both alternative species and rodents. The distinct advantages of these approaches are that they are faster and less expensive than traditional data generation approaches while retaining the ability to assess potential toxicity in systems that have intact biological complexity, architecture, and metabolic

---

the word “toxicokinetics” (TK) to distinguish the chemical as a toxicant. Both terms are in common use, and might appear in the text of this document. They relate to the same processes and are interchangeable.

capacity (i.e., compared with *in vitro* systems). Alternative species have additional advantages: toxicity can be observed over the lifespan of the organism (nonmammalian species lifespans are shorter compared to mammalian species); adverse health effects can sometimes be more directly observed and provide context for interpreting molecular data; and complex effects, such as birth defects and neurobehavioral deficits, can be studied in intact biological systems. The development of the short-duration, *in vivo* study Tier 2 prototypes led us to the following inferences:

- Perturbations in molecular mechanisms are useful indicators of potential toxicity, but the correlation of the specific adverse effects across species for similar doses and routes of exposure can be complicated. This is true for traditional as well as new data types.
- Species extrapolation issues in the use of nonhuman species remain a challenge when interpreting the relevance of the effects for humans, and in estimating equivalent doses for a given response level. Many important molecular mechanisms and biological processes are well conserved across species, but the relationships of molecular events to specific downstream adverse outcomes, in some instances, have drifted with evolution. Thus, the specific outcome of a perturbed molecular mechanism might differ among species.
- Alternative species data (e.g., zebrafish developmental assay data) are suitable for identifying hazards and evaluating the potency of chemicals to cause an adverse effect, and could be used for screening and prioritization. Augmented with sufficient supporting data (e.g., AOP information) and exposure concentrations that are relevant to human exposures, these data could be suitable for determining toxicity values for limited decision-making.
- The prototype example based on the short-term *in vivo* rodent study presents transcriptomic data that correlated well with dose-response relationships based on traditional cancer and noncancer endpoints. In this example, the observed transcriptomic events were nonspecific relative to the observed adverse effects in the traditional studies, that is, transcriptomic events did not predict a specific adverse outcome. The logic, however, is that toxicity must be preceded by changes in gene expression and, hence, the concentration at which gene expression changes occur could be used in prioritization and screening, and in determining a BMD. Due to the associative versus causal nature of these studies and uncertainties in the predictive nature of transcriptomic events for adverse effects, these data are considered suggestive. As with other approaches, supporting data—such as mechanistic information, consistent results across multiple studies, and experimental interventions that demonstrate causal relationships—would increase the confidence in the overall evidence and expand use of these data in risk assessment. Although current experience with this model is limited, wider use in the future is anticipated.
- In addition to extrapolation issues for molecular events, differences in the toxicokinetics (i.e., the absorption, distribution, metabolism, and excretion) of chemicals among species might exist, confounding comparison of target tissue levels and responses.

In summary, these Tier 2 approaches are promising, will be deployed in the near future particularly for screening and prioritization, and will be further evaluated for use in estimating toxicity values.

3.3 Tier 1: Screening and Prioritization

This section summarizes *in vitro* new high-throughput (HT) and high-content (HC) approaches available to develop data for screening and prioritizing large numbers of chemicals (i.e., tens of thousands of chemicals) into categories for focused research, further testing, or further assessment. Tier 1 chemical rankings rely on QSAR models, HTS/HCS assay data, statistical correlations between *in vitro* and *in vivo* assays, and systems models that focus on molecular targets and chemical perturbations to susceptible biological pathways that might result in adverse effects or clinical disease. The Tier 1 prioritization and screening prototypes demonstrated use of a variety of these new approaches including the following:

- QSAR models, read-across, and high-throughput virtual molecular docking (HTVMD) models (discussed in Section 3.3.1);
- High-throughput screening (HTS)/high-content screening (HCS) assays and the ToxCast Program (discussed in Section 3.3.2);

Section 3.3.3 discusses the risk assessment implications across the Tier 1 prototypes.

HT *in vitro* assays are used to probe MIEs (e.g., activation or inactivation of specific receptors, enzymes, or transport proteins) or other key events in these biological pathways, and to replace more costly and time-consuming assays to assess the potential for adverse outcomes. Correctly identifying the linkages from a MIE to an assay endpoint to potential for adversity is key to the relevance of each assay for use in Tier 1 risk assessments. The critical support for this linkage comes from statistical modeling using *in vivo* and *in vitro* data on the same chemicals, from literature reviews and biological expertise, and from adequately developed systems biology models with acceptable simulations of normal and altered biological processes (e.g., virtual tissue models or other types of systems biology models). The knowledge gained in developing Tier 2 and 3 assessments will provide further context and support continual improvement in interpreting Tier 1 data.

EPA's Chemical Safety for Sustainability (CSS) program is actively researching systems approaches and developing tools to understand links between exposures to chemicals and disruptions in pathways that lead to disease (EPA 2012b). The CSS research aims to increase the efficiency and speed of chemical evaluations dramatically, and to support assessment of potential effects from chemical exposure at critical lifestages (the embryo and childhood), and on susceptible populations associated with factors such as genetic differences or coexisting diseases. The program within CSS that focuses on developing automated chemical screening technologies is the Toxicity Forecaster (ToxCast™) program. ToxCast is a multiyear effort launched in 2007 to evaluate HTS/HCS assays with living cells, or isolated cellular components (e.g., proteins, nuclear receptors, transcription factors, enzymes), and to develop rapid approaches for assessing adverse health effects of chemicals. A key goal of the ToxCast program is to protect human health by identifying chemicals that are of potential concern and require additional testing, and to limit the number of animal-based toxicity tests. The main thrust of the research aims at understanding correlations and

linkages between molecular/cellular perturbations and apical toxicity endpoints (adverse outcomes). To resolve these linkages, statistical and computational (*in silico*) models are being developed that compare the chemical effects results from ToxCast *in vitro* assays to the adverse effects outcomes from thousands of *in vivo* animal toxicity studies on hundreds of chemicals that have been compiled in EPA's Toxicity Reference Database (ToxRefDB). EPA's ToxCast assay results and databases are freely available to the science community and to the public (EPA 2014l).

NIH is also developing and deploying several tools and databases to evaluate assays, pathways, and underlying mechanisms. These include the Tox21[38] data, BioSystems (NCBI 2014a; database of mechanistic networks), BioAssay Research Database (NCBI 2014e; data on more than 35 million compounds and thousands of assays and experiments), and the 1000 Genomes Browser.[39] A major thrust of these NIH resources is to summarize, and make publicly available, information found in the scientific literature, thus facilitating transparent meta-analyses of data and broad acceptance of approaches within the scientific community.

HT/HC assays and models have several purposes for risk assessment. One is to generate data for Tier 1 assessments that screen and prioritize chemicals for potential toxicity. Prioritization identifies the subset of tested chemicals that could disrupt normal pathways or bind to critical targets to elicit adverse outcomes. Alternatively, prioritization could identify chemicals that are of less concern for toxicological effects. The *in vivo* dose leading to the concentration in the *in vitro* assays needed to activate targets or perturb networks can be estimated with reverse toxicokinetic models (see Section 3.3.2.3). These estimates are then interpreted within the context of real or potential environmental exposures, duration and frequency of those exposures, and other relevant information to rank tested chemicals for level of concern, or to identify subsets to advance to Tier 2 or Tier 3 testing or further evaluation. In some cases, the data developed in Tier 1 could be used to supplement the evidence for reference values derived in Tiers 2 and 3 assessments, especially with respect to identifying AOPs and AOP networks associated with chemical-induced disease (see Table 10). HT methods also might be used for rapid data generation to help risk assessors and managers make urgent decisions. Examples of decision-making where high-throughput prioritization and screening are useful include emergency response or urgent, need-to-identify chemicals of potential concern. Specific examples include the HT-based evaluations of dispersants

[38]Several federal agencies collaborate in the Tox21 program: Environmental Protection Agency, National Institute of Environmental Health Science/National Toxicology Program, the NIH Center for Advancing Translational Science/National Chemical Genomic Center, and the Food and Drug Administration. Collaborators conduct screening with many of the same aims as EPA's ToxCast program but cover more chemicals with fewer HTS assay technologies (Tice et al. 2013).

[39]The 1000 Genomes Browser is an interactive graphical viewer that enables users to explore variant calls, genotype calls, and supporting evidence (such as aligned sequence reads) that have been produced by the 1000 Genomes Project. The project is an international collaboration to produce an extensive public catalog of human genetic variation (including SNPs and structural variants, and their haplotype contexts) to support GWAS and research on human genetic variants and their contribution to disease. The genomes of about 2500 unidentified people from about 25 populations around the world will be sequenced (1092 have been sequenced to date). The results are freely and publicly accessible to researchers worldwide.

used in the Deep Water Horizon Gulf oil spill (Judson et al. 2010); the ongoing prioritization and screening of potential endocrine disruptors as part EPA's effort under the Food Quality Protection Act (limitations on pesticides in food) and the Safe Drinking Water Act (EPA 2011c); and in research applications (Knudsen et al. 2011a).

Table 10. Summary of Tier 1 NexGen Approaches, Including Strengths, and Weaknesses

TIER 1: SCREENING AND PRIORITIZATION PROTOTYPES		
	New QSAR Models	**High throughput *In Vitro* Assays**
Approaches:	• Uses structural characteristics and experimental data from chemical analogs to predict toxicity for various endpoints, metabolism, fate, and chemical groupings (based on similarity) for data-poor chemicals	• Experimentally measures concentration-dependent, chemically induced alterations in biological functions using range of specific and sensitive *in vitro* assays • Infers potential adverse outcomes based on existing knowledge of other chemicals and potential importance of selected biological processes
Strengths:	• Chemicals can be classified rapidly and inexpensively • Some QSAR models can generate quantitative values such as the LOAEL that can be used both to rank chemicals and as a point of departure for reference value derivations (albeit with considerable uncertainty associated with that value) (OECD 2014e) • SAR models provide quality characterizations useful for binning chemicals into groups for read-across	• Rapid, relatively inexpensive, multiple bioassay options available • Research on key pathways and new assays rapidly progressing, including alternative test species assays that improve the representation of *in vivo* responses • Systems biology models continue to evolve with increasing amounts of knowledge and increases in their predictive utility and context for interpreting the *in vitro* results
Weaknesses:	• If the physical chemistry or structures of chemicals being evaluated differ significantly from the chemicals used to develop the models (the training set) or have fragments not represented in the training set, results likely to be unreliable • Active metabolites are not represented in the results for parent compounds • Major issues exist around characterizing the uncertainty in QSAR and related read-across approaches, and in the transparency of some models (see Ball et al. 2014; OECD 2004b; Patlewicz et al. 2013a)].	• Coverage of important biological processes is incomplete, cell lines generally not metabolically competent and vary widely from their *in situ* counterparts, interactions among cell types or tissues cannot be evaluated in *in vitro* assays • Volatile and chemical gases cannot currently be tested • Systems biology models (and approaches) require consistent support and iterative laboratory collaborations to improve and update the models continually (i.e., short-term planning is inadequate)

3.3.1 QSAR Models, Read-across, High-throughput Virtual Molecular Docking (HTVMD) Models

QSAR models are regression or pattern recognition models used in risk assessment to classify or predict target toxicities, chemical potency, exposure potential, and the like, as a function of one or more chemical descriptors. The descriptors are generally inherent physiochemical properties of the chemical, such as atomic composition, structure, substructures, hydrophobicity, surface area charge, and molecular volume. QSAR models correlate inherent properties of the two-dimensional or three-dimensional chemical structure of an unknown chemical, the "query" chemical (as input

parameters in the model run), with similar properties for a set of chemicals having known toxicological or exposure potential called the "training set" (EC 2010; EPA 2014g; Goldsmith et al. 2012; OECD 2014c; Venkatapathy and Wang 2013; Wang, N. et al. 2012c). QSAR models are run on high-speed computers, and the output is thus considerably less costly and orders of magnitude faster than *in vitro* or *in vivo* assays. Interpreting QSAR results for use in hazard and dose-response assessment, however, requires expertise, and issues exist with transparency and uncertainty characterization.

A variety of QSAR models and support tools are available to choose from (Hansen et al. 2011; JRC 2014; OECD 2014d). Each model has its own set of assumptions and chemical domains of applicability. Although QSAR is not a new technique, what is new are more concerted efforts to validate predictive accuracy relative to authoritative, traditional based toxicity values (Golbraikh et al. 2012; Venkatapathy et al. 2004; Wang, N. et al. 2012b; Wang, N. et al. 2012c; Wang, N. et al. 2011).

QSAR models have been used most commonly in the classification of data-poor chemicals with unknown hazard or exposure potential. Each model can generate quantitative estimates for various outcomes, for example, kinetic parameter values, a rodent oral or inhalation LD_{50}, a fish LC_{50}, or a rodent maximum tolerated dose. The commercially available TOPKAT model (TOPKAT 2014) is the only QSAR model (at the time of this report) that generates a rodent quantitative lowest observable adverse effect level (LOAEL) and, importantly, has been evaluated in studies published in the peer-reviewed literature (Venkatapathy et al. 2004; Venkatapathy and Wang 2013). The TOPKAT generated LOAEL can be used to rank chemicals and as a point of departure (POD) to compare with existing reference values, albeit with a considerable number of caveats concerning confidence in those QSAR-based POD values. Significant limitations in the TOPKAT model include a database in need of updating with new information since 2004, and a lack of transparency.

Structure-activity relationship (SAR) models use a similar modeling approach but generate only qualitative characterizations. Qualitative characterizations can be used to categorize chemicals for specific hazards (e.g., suspected carcinogen, likely mutagen, potential developmental toxin). An HT SAR approach popular in the European Union is called read-across. Substances with physicochemical and human health or ecotoxicological properties or environmental fate properties that are similar, or that follow a regular pattern (usually because of structural similarities), can be considered as a "group of substances." These groups of chemicals are used to predict the physicochemical properties, human health effects, or environmental effects of a new "target substance(s)" that has inherent properties similar to those of the groups. The predictions are made by interpolating to other substances in the group called "reference substance(s)" that have had adequate testing for these characteristics, and become "source substance(s)" for read-across (OECD 2014b, d). The term analog approach is used when read-across is employed within a group of a very limited number of substances for which trends are not apparent. The simplest case is read-across from a single source substance to a target substance. When a group contains more substances, the term category approach is used (see Box 8).

The grouping of substances and read-across offer a possibility for meeting the standard information requirements of the European Union's REACH[40] regulation (requirements set in Annex XI, 1.5) (REACH 2014). The read-across approach must be considered on an endpoint-by-endpoint basis due to the different complexities (e.g., key parameters, biological targets) of each endpoint. If the read-across approach is adequate, unnecessary testing can be avoided.

Box 8. From: Use of Category Approaches, Read across and (Q)SAR: General Considerations. ECETOC Technical Report 116 (Patlewicz et al. 2013a)

There are many endpoints where read across can be applied and these range in complexity and sophistication from simple physiochemical and acute/local effect to repeated dose/systemic and reproductive toxicity. This range of endpoints translates to a range of complexity of approaches that needs to be developed, i.e., not as simple as one size fits all. The foundation of many categories of read across justification is that the substances are similar in structure (same functional groups) or have a common/shared metabolic pathway or precursor. … Data on toxicokinetics can be a key piece of evidence to support these justifications."

The Organization for Economic Cooperation and Development (OECD) provides a free downloadable QSAR software package, the QSAR Toolbox, that is intended for use by governments, the chemical industry, and other stakeholders to assess potential human and ecological chemical toxicities for data-poor chemicals (OECD 2014c). The QSAR Toolbox estimates the potential toxicity of a compound of interest based on the available information for structurally similar analogs, and uses similarities or trend analysis to construct the categories of chemicals for read-across screening purposes even if only a few of the members in the category have available test data. Read-across has become one of the most widely used approaches under REACH (Patlewicz et al. 2013b). The method's popularity is driven not only by its relative simplicity and the online availability of the QSAR Toolbox (ECETOC 2013; ECHA 2012; OECD 2014d), but also because it provides some information to evaluate chemicals of interest when no other information is available.

OECD and others have developed guidance for use of QSAR models for regulatory purposes (NAFTA 2012; OECD 2004b). Documenting and addressing uncertainty in the read-across results, however, remains a major challenge. Within the European Union, ECHA is developing a framework to facilitate a more transparent and structured approach to identifying and assessing uncertainty associated with the use of read-across (Ball et al. 2014; Patlewicz et al. 2013a). Others in the industrial sector are also developing approaches to address uncertainty systematically in read-across results (Blackburn and Stuard 2014).

At EPA, QSAR models are being used to screen, rank, and categorize chemicals for level of concern in a variety of EPA programs, including Superfund mitigation; the Office of Chemical Safety and

[40]**REACH – Registration, Evaluation, Authorisation and Restriction of Chemicals**. REACH is a regulation of the, adopted to improve the protection of human health and the environment from the risks that chemicals can pose, while enhancing the competitiveness of the European Union chemicals industry. It also promotes alternative methods for the hazard assessment of substances to reduce the number of tests on animals. REACH requirements became effective June 1, 2007 and are implemented by the European Chemicals Agency (ECHA 2014).

Pollution Prevention High Production Volume Challenge Program and Pre-Manufacture Notice review process; the Office of Chemical Safety and Pollution Prevention/Office of Water Endocrine Disruptors Screening Program (Weiss et al. 2012); and the Office of Water Candidate Contaminant List. The QSAR models used by EPA include the Sustainable Futures Initiative suite of models, the OECD QSAR toolbox models (OECD 2004b, 2014c), HTVMD (Rabinowitz et al. 2008), MetaCore (Teschendorff and Widschwendter 2012; van Leeuwen et al. 2011), and the TOPKAT model (Rakyan et al. 2011; Venkatapathy et al. 2004).

High-throughput virtual molecular docking (HTVMD) models use a ligand-based chemoinformatics strategy to predict relationships between various attributes of ligands and their binding to known targets. These models, which are increasingly being used in risk assessment, can screen thousands of chemicals for the potential affinity of their three-dimensional structures to bind to active protein binding sites. HTVMD models have been used in the pharmaceutical industry for years to identify candidate drugs. These models also can be used to estimate the likelihood that a chemical of toxicological interest would bind to a target protein, for example, the potential affinity of a chemical as a direct agonist of the estrogen receptor.

Limitations in current HTS/HCS assays include difficulties in evaluating the toxicity of metabolites, volatiles, and limited solubility compounds such as metals. QSAR and HTVMD models can provide some information to address these gaps in chemical coverage. Recent advances in high-performance computing support simultaneous runs of QSAR and HTVMD models, dramatically decreasing the time to discovery. The U.S. Army Medical Research and Materiel Command, for example, has recently published their version of a Docking-based Virtual Screening pipeline that facilitates the use of the AutoDock molecular docking software on high-performance computing systems (Jiang et al. 2008).

Results from these rapid, computationally based methods (e.g., QSAR, read-across, molecular docking models) can add to the evidence in assessments based on more traditional data (when available) and advance the speed and accuracy of chemical screening (Golbraikh et al. 2012; Lock et al. 2012; Rusyn et al. 2012; Sedykh et al. 2011; Wignall et al. 2012). Continued improvements and transparency in these models, and the criteria for interpreting the data, are anticipated to support their use for chemical screening and prioritization, and in the design of new products and chemical processes that minimize harm to health and the environment (i.e., green chemistry).

3.3.2 High-throughput and High-content (HTS/HCS) Screening Assays

High-throughput screening (HTS) and high-content screening (HCS) assays are major tools used for early evaluation of chemicals and to determine the chemicals' ability to perturb molecular pathways (Judson et al. 2013; Judson et al. 2011; Kavlock et al. 2012; Sipes et al. 2013; Tice et al. 2013). For example, as part of EPA's ToxCast program, the following (EPA 2014l) were conducted:

- A chemical prioritization project compiled and analyzed data on 1877 chemicals, including pesticides; food, cosmetics, and personal care ingredients; pharmaceuticals; and industrial chemicals.

- HT testing used a battery of 782 *in vitro* assays across 7 distinct technologies and multiple biological formats (cell-free, cell lines, and primary cells from multiple tissue types).
- All 1877 chemicals were tested in a subset of 185 endocrine-related assays for nuclear receptors, steroidogenesis, and CYP P450 assays.

Several predictive models are undergoing further development (see Box 9). Much of the HTS/HCS methodology was developed to aid the pharmaceutical and biotechnology industries in the drug discovery process, where screens are needed for millions of candidate compounds to identify candidate drugs for a target of interest (e.g., a receptor or enzyme) (Bleicher et al. 2003; Mayr and Bojanic 2009). The technology has broader use in approaches previously called chemical genetics (or sometimes, chemical biology), in which small-molecule screening is used to identify probes for biological signaling networks and cellular phenotypes (Schreiber, S. 2003). More recently, toxicologists have become interested in these assays because targets of pharmaceuticals and research chemicals might be similar to those involved in disease processes induced by exposures to environmental chemicals (Houck and Kavlock 2008). A large data matrix of toxic chemicals versus appropriate HTS assay results provided toxicologists a novel and promising approach for identifying AOP networks leading to adverse phenotypic changes.

Box 9. Examples of Current Research on Predictive High Throughput and Content Models

Endpoints

- Liver tumors: Judson et al. (2010)
- Hepatocarcinogenesis: Shah et al. (2011)
- Rat fertility: Martin et al. (2011)
- Rat rabbit prenatal developmental tox: Sipes et al. (2011a)
- Zebrafish development: Sipes et al. (2011b)

Pathways

- Endocrine disruption: Reif et al. (2010)
- Microdosimetry: Wambaugh and Shah (2010)
- Differentiation: Chandler et al. (2011)
- Angiogenesis: Kleinstreuer et al. (2011a)
- Cancer hallmarks: Kleinstreuer et al. (2013b)
- Endocrine activity: Rotroff et al. (2012)

The underlying technologies for HTS assays are well known, and the discussion here focuses on a broad description of the types of assays and some of the key issues to be considered when designing *in vitro* assays for Tier 1 assessments. HTS assays can be divided broadly into two types: cell free/biochemical assays and cell-based assays. Cell-free assays typically test for the direct interaction of a test chemical with a specific protein such as a receptor, enzyme, or transcription factor. Measures of interaction include activation, repression, or inhibition of the protein's activity. In cell-based assays, a cellular readout can be molecular based (e.g., changes in gene or protein expression) or phenotypic (e.g., cytotoxicity, changes in cell morphology). The selection of the cell system is critical for cell-based assays. These assays have been developed using a variety of primary cell types from various organs and species, immortalized cell lines, or stem cells (Dick et al. 2010; EPA 2014k; NCBI 2014e). Each type has strengths and weaknesses. For example, immortalized cell lines generally produce very reproducible screening results over long periods of time due to the continuous growth and stability of the cell lines. The disadvantage is the significant differences in these cells from their normal (i.e., nonimmortalized) *in vivo* counterparts with respect to the completeness or representation of physiological processes. These differences might result in different outcomes when subject to comparable chemical exposures. The converse holds true for most primary cells, that is, they better represent normal physiological responses, yet are more challenging with respect to consistent, reproducible screening results. Co-culture systems combine different cells in an attempt to mimic

in vivo systems and their complex cell-cell signaling networks (Berg et al. 2010). Advanced culture methods compatible with HTS also are being developed, for example, three-dimensional collagen matrices designed to enhance maturation of induced pluripotent stem cell-derived hepatocytes (Gieseck et al. 2014). These systems improve the consistency and longevity of the test cell population (compared with primary culture cells) and provide better representation of normal biology (relative to immortalized cells). Certain whole organisms, including *Caenorhabditis elegans* and zebrafish embryos, are providing valuable new HTS assay data (Kanungo et al. 2014; Parng et al. 2002; Smith, M. V. et al. 2009).

3.3.2.1 The First-Generation of Predictive Models from ToxCast

The following discussion summarizes the main results of work conducted in EPA's ToxCast program focusing largely on the published work from Phase I of ToxCast, which tested about 300 chemicals, primarily active ingredients in pesticides (Knudsen et al. 2011a; Rotroff et al. 2013; Sipes et al. 2013). The Phase II results, which extend testing to as many as 1877 chemicals, have just recently been released (EPA 2014j).

Models for Reproductive, Developmental, Chronic, and Cancer Endpoints

Several first-generation (Phase I) models have been published to date, including ones for reproductive, developmental, and chronic/cancer endpoints (Judson et al. 2008; Knudsen and Kleinstreuer 2011; Martin et al. 2012; Martin et al. 2011; Martin et al. 2009b; Sipes et al. 2011a; Wetmore et al. 2013; Zaldívar et al. 2012). These models are being tested and refined using the newest (Phase II) ToxCast data. An important point about these models is that the *in vitro* data are principally derived from human cells, while the *in vivo* data are from rodents and rabbits. The following text on the models for reproductive toxicity, developmental toxicity, developmental vascular disruption, and cancer is reproduced from Judson et al. (2014).

> *"Reproductive Toxicity Model*: Initial models of reproductive toxicity were built using the data set compiled by Martin et al. (2009a). This data set compiled information on 75 reproductive effects for 256 chemicals with data from both ToxCast and guideline studies on multigeneration rat reproductive guideline studies performed as part of pesticidal active ingredient registrations. A total of 19 parental, offspring or reproductive endpoints had a sufficiently high incidence after chemical exposure and were used as predictive end-points in the model. These included reproductive performance indices, male and female reproductive organ pathologies, offspring viability, growth and maturation, and parental systemic toxicities. Next, these end-points were combined with the ToxCast data to build a model of generalized reproductive toxicity. A reproductive toxicant was defined as a chemical with a reproductive adverse effect seen at <500 mg/kg/day. A total of 68 chemicals in the data set were considered reproductive toxicants. Using the *in vitro* assay data from ToxCast, a linear discriminant analysis (LDA) model was constructed that predicted the reproductive toxicity with a 74 percent balanced accuracy (BA = mean of sensitivity and specificity) based on cross-validation and a 76 percent BA using an external validation set. The *in vitro* assays used in the model included activity in nuclear receptors (estrogen receptor, androgen receptor, peroxisome- proliferator-

activated receptor [PPAR]), cytochrome P450s, G-protein-coupled receptors, and other cell signaling pathways. This model was also evaluated for its utility in prioritizing chemicals for further testing based on a scenario where many chemicals were tested *in vitro*, but where only a few could be tested *in vivo* because of cost and animal welfare considerations (Martin et al. 2011). Two regulatory environments were evaluated in this study—one consistent with industrial chemicals where little data are required to be generated unless there is prior evidence of risk (screen in) and another where many studies are required for registration, but the U.S. EPA has the ability to waive (screen out) certain studies.

Developmental Toxicity Model: Models of prenatal developmental toxicity used data compiled from ToxRefDB on guideline rat and rabbit developmental toxicity studies (Knudsen et al. 2009). A total of 383 rat and 368 rabbit studies were available, covering 387 chemicals, mostly pesticidal active ingredients. Of these chemicals, 283 were tested in both species, and, of those, 53 chemicals were specifically developmentally toxic (no overt maternal toxicity or maternal toxicity at doses higher than observed for the developmental defects). The primary expressions of developmental toxicity in pregnant rats were fetal weight reduction, skeletal variations and abnormalities, and fetal urogenital defects. Relative to rats, general pregnancy/fetal losses were more frequently observed in the rabbit as were structural malformations to the visceral body wall and CNS [central nervous system]. Species-specific models were built on these data, linking *in vitro* ToxCast data to developmental defects (LDA with cross validation) (Martin et al. 2012). Specifically, 271 chemicals (187 unique) with ToxCast and ToxRefDB data were used, with 251 for the rat model (146 identified as developmental toxicants) and 234 for the rabbit model (106 identified as developmental toxicants). A developmental toxicant was defined as eliciting any significant end-point (i.e., fetal weight reduction, various malformations, prenatal loss) regardless of the maternal toxicity dose. The overall risk of a chemical causing developmental defects was linked to disruption of the following main targets and pathways: transforming growth factor beta (TGFβ), retinoic acid receptor (RAR), and G-protein-coupled receptors in rat; and interleukins and chemokines in rabbit. Species-specific models had a BA of about 70 percent. A key finding was that the molecular effects driving prenatal developmental toxicity showed strong species dependence in prediction models for pregnant rats and rabbits. Because the same set of *in vitro* assays was used for both species models, the differences are assumed to reflect model input parameters related to (i) the chemical space tested in each species; and (ii) the apical end-points (*in vivo* outcomes) recorded for each species, toxicokinetic differences between rats and rabbits, and/or toxicodynamic differences between the responses in pregnant dams and their concept uses for either species.

Developmental Vascular Disruption Model: Several of the molecular targets associated with developmental defects suggested a broad linkage between disruption of vascular development and the emergence of gross phenotypic

developmental defects. This hypothesis led to the concept of 'putative vascular disrupting compounds' (pVDCs) (Hanahan and Weinberg 2000; Kleinstreuer et al. 2011a; Knudsen et al. 2009; Sipes et al. 2011a). An AOP linking multiple molecular initiating events to outcomes was developed around the biomedical literature and Mouse Genome Informatics (MGI) database to provide a framework for identifying pVDCs based on ToxCast *in vitro* signatures. Particular targets included inflammatory chemokine signaling (CK), the vascular endothelial growth factor (VEGF) pathway, and the plasminogen-activating system (uPAR). Consistent with the species dependence of prediction models built for prenatal developmental toxicity in pregnant rats and rabbits (Martin et al. 2012), we also observed species differences in models predicting pathway-level sensitivity to angiogenic signals, particularly those mediated by CK and uPAR pathways. This suggests a mechanistic link to species-dependent processes for inflammatory responses and extracellular (ECM) remodeling, respectively. The group of pVDCs with rat developmental toxicity correlated with down-regulation of pro-inflammatory CK assays, whereas pVDCs with rabbit activity often resulted in up-regulation of these signals. The rabbit pVDCs generally showed greater bioactivity across assays, which can be inferred to entail ECM degradation and release of angiogenic growth factors. The observed *in vivo* developmental toxicity also showed a distinct trend across species, with skeletal malformation in rats and prenatal death in rabbits being the most prevalent end-points for the pVDCs (Sipes et al. 2011a). To further investigate this linkage, a cell/tissue-level dynamic signaling *in silico* model was developed (Kleinstreuer et al. 2011a) using the CompuCell3D (CC3D) software (Swat et al. 2012). The *in silico* model could recapitulate self-directed assembly of endothelial cells into a completed vascular network utilizing signal-response pathways involving an exchange of CK, VEGF, and uPAR among several cell types. By incorporating parameters from ToxCast HTS data into this 'virtual tissue model', the concentration-dependent disruption of angiogenesis was shown for 5HPP-33, an anti-angiogenic thalidomide analog.

Cancer Model: We also have published a first-generation prediction model linking *in vitro* effects and the likelihood that a chemical will be an *in vivo* carcinogen (Judson et al. 2008). This model began with the hypothesis that chemicals perturbing cancer hallmark processes would increase the likelihood of those chemicals being carcinogens (Hanahan and Weinberg 2011; Thomas, R. S. et al. 2013c). To test this hypothesis, univariate associations were calculated between each gene tested by the ToxCast assays and each cancer end-point (rat or mouse) in ToxRefDB. We found that the vast majority of cancer-linked genes (defined as having an odds-ratio > 2, with confidence intervals not overlapping with zero after permutation testing) were either hallmark-associated or involved in xenobiotic metabolism. A scoring function was used that combined the cancer-associated gene hits for each chemical into an overall score. This was applied to an external test set of 33 chemicals that were not used in the model development process. The results were that the model distinguished between carcinogens and noncarcinogens with statistical significance

(p = 0.024). Future work on all of models will expand them to look in more detail at the molecular mechanisms linked to the adverse outcomes and to forward validate using ToxCast Phase II data."

3.3.2.2 Summary of ToxCast Phase I Results (This section is reproduced from Judson et al. 2014)

[Note: This section is reproduced from Judson et al. 2014]

"The goals of [ToxCast] Phase I largely have been met, and include the following *demonstrating*: (i) that a large set of environmentally relevant chemicals can be screened in a diverse battery of *in vitro* assays; (ii) that predictive models of toxicity can be developed using these data; and (iii) that *in vitro* pharmacokinetic data can be integrated with the *in vitro* assay data, enabling us to make initial quantitative comparisons with *in vivo* rodent toxicity data. That said, a number of challenges lie ahead. Some of these have been outlined by other researchers who performed independent analyses of the ToxCast data (Benigni 2013; Sonich-Mullin et al. 2001). One challenge is presented by the broad diversity of chemicals, chemical-biological activities *in vitro* and chemical effects *in vivo*. At the very least, these pose a classic statistical power issue. For instance, if there are N different mechanisms by which a chemical can cause a specific phenotype, one needs a significant multiple of N examples for each such pathway-end-point pair in the data set to discover this linkage using purely statistical methods (Knudsen and Kleinstreuer 2011). This argues for the need to increase the size of the data set (number of chemicals)*, and data are now available from* Phase II of ToxCast" (*italics* are Editor's revised text).

The amount of high-quality *in vivo* toxicity data will increase much more slowly than the amount of HTS/HCS data, hampering the development of predictive models based solely on statistical analysis. Most chemicals with existing traditional data have been captured in ToxCast and Tox21. In addition, although tremendous progress is being made in understanding the network of events that underlie disease and in developing assays to test for these events, the field is still in its infancy. In particular, HT assays generally measure changes in important key events or processes, rather than an integrated indicator of adverse outcomes. Variability or confounding factors in *in vivo* conditions (e.g., species, tissue, lifestage, metabolism, complex interactions), some of which might be difficult to evaluate in *in vitro* systems, might lessen the utility of HT approaches in predicting disease. As a consequence, characterizing the results of HT testing as indicative of alterations in biological processes rather than as predicting disease (Thomas, R. S. et al. 2012b) is generally more reasonable. As discussed above, more complex, cellular and multiscale, biologically based models are therefore needed to interpret HT data, and to simulate outcomes indicative of multiple levels of biological organization and interactions. Such models could leverage and incorporate biological knowledge and expertise on the etiology of disease. EPA's VT modeling research continues to progress toward that end.

Further advances are needed in developing HT quantitative reference values for use in risk assessment. One approach is to use HT/HC *in vitro* and *in vivo* data to develop reference values that support or supplement traditional values that require extensive *in vivo* animal test data. Points of departure derived from HT/HC data might be used to guide further testing, and for many chemicals

preclude the need for specific hazard identification. This approach could be viable because expected exposures are likely much lower than the *in vitro*-derived points of departure (i.e., acceptable margins of exposure from a risk management perspective) (Rotroff et al. 2010; Wambaugh et al. 2013). The goal would be to have a significantly smaller set of chemicals for which more *in vivo* data intensive reference values, and follow-up testing, would be needed.

Although these challenges are daunting, *in vitro* methods and computational models have already demonstrated proof-of-concept that HTS/HCS assays likely will improve risk predictions for human and ecosystem health for thousands of currently untested chemicals, as data increase and methods evolve.

3.3.2.3 Toxicokinetics

Toxicokinetic models have been developed to extrapolate the concentration of a chemical that is used in the HTS assays to an equivalent dose that would be delivered to a target in a test animal or human, providing key information for use in dose-response characterization. As previously mentioned, the HTS assays are run in concentration-response format. The potency of each chemical in each assay can be summarized using AC_{50} or lowest effective concentration values, depending on the type of dose-response data collected. The potency values among the *in vitro* assays, along with other chemical information, have been proposed for use in hazard identification (Martin et al. 2011; Sipes et al. 2011b) and prioritization of chemicals for further testing (Reif et al. 2010). The relationship between the *in vitro* concentration of the chemical in the well to the concentration of the chemical in the blood or target tissue (*in vivo*), however, can be complex and can depend on variables not captured in the HTS assays. These variables include bioavailability, clearance, and protein binding (Wetmore et al. 2012).

In vitro-to-*in vivo* extrapolation (IVIVE) is a process that uses data generated within *in vitro* assays to estimate *in vivo* drug or chemical fate. In the past, IVIVE has been developed and applied in the pharmaceutical industry predominantly to estimate therapeutic blood concentrations for specific candidate drugs and to identify potential drug-drug interactions (Chen, Y. et al. 2012; Gibson and Rostami-Hodjegan 2007; Shaffer et al. 2012). Due to both legislative mandates and public pressure for increased information on potential chemical toxicity, IVIVE is increasingly being used to predict the *in vivo* toxicokinetic behavior of environmental and industrial chemicals (Basketter et al. 2012).

Reverse dosimetry uses a PK model to determine a plausible exposure concentration based on a measured or estimated internal concentration of a chemical at a target site (or based on a surrogate internal metric such as a biomarker of exposure). At the population level, probabilistic reverse dosimetry uses a distribution of internal concentrations to identify the most likely exposure concentrations (or intake doses) experienced by a population of interest (Grulke et al. 2013). A combination of IVIVE and reverse dosimetry can be used to estimate the daily human oral dose (called the oral equivalent dose) necessary to produce steady-state *in vivo* blood concentrations that are considered equivalent (with respect to chemical concentration at potential targets) to the dose delivered *in vitro* at the AC_{50} or lowest effective concentration values. The estimated *in vivo* exposures likely to produce adverse effects based on *in vitro* data can be generated for each assay across the more than 600 *in vitro* assays (Rotroff et al. 2010; Wetmore et al. 2012). These estimates

of potentially adverse exposure levels can be compared with model estimates of actual exposures for chemicals based on production volume or use patterns (Mitchell et al. 2013; Wambaugh et al. 2013; Wambaugh and Shah 2010).

3.3.2.4 Virtual Tissue Modeling

A major challenge in the use of *in vitro* data is how best to organize and interpret the information for relevance to *in vivo* human responses. Purely statistical methods that treat the data as just a set of numbers with no biological context have inherent limitations, including uncertainty about biological relevance, and an increase in chance correlations when correlating large numbers of explanatory variables to only one endpoint (Benigni 2013; Thomas, R. S. et al. 2012b). One approach to address this challenge is to develop network models of AOPs (Ankley et al. 2010; Boobis et al. 2008; Kleinstreuer et al. 2011b; Meek et al. 2003; Seed et al. 2005; Thomas, R. S. et al. 2013c). These network models are essentially hypotheses constructed from knowledge and data about the biological processes. Proposed AOP networks provide additional biological context to interpret the *in vitro* assay results and statistical analyses but they do not address the multiple testing issues inherent in the statistical approach. A step further is to develop and use more complex, cellular and multiscale, biologically based models (often referred to as VT models). VT models incorporate knowledge of the structure of the biological pathways being altered (including PK information), and explicitly address and represent the spatial and temporal dynamics of multiple levels of biological organization (DeWoskin et al. 2014; Knudsen and Daston 2010).

VT models provide an experimental and theoretical framework for the systematic and integrative analysis of complex multicellular systems. They capture the flow of molecular information across cellular and biological networks, and process this information computationally into higher order responses that ideally simulate a potential adverse outcome. Responses to perturbation depend on network topology, system state dynamics, and collective cellular behavior. For agent-based VT models, these simulations are enabled from individual cellular behaviors in a multicellular field that can result in emergent properties. Emergent properties are behaviors that arise from interactions of parts at the next higher level of a system (e.g., functions, phenotypes) that are not apparent from knowledge about the behavior of the parts alone. VT models address both the relevance and multiple comparison issues by prioritizing the most relevant assays and interpreting their results in a systems biology context and are the focus of EPA's VT modeling research. The initial focus is to develop virtual embryo (v-Embryo™) models for various developmental effects and the virtual liver (v-Liver™) for hepatotoxic effects.

The goal of the v-Embryo project is to provide a rapid, hypothesis- and chemical-testing platform capable of estimating the probability of adverse effects on the developing embryo from exposure to environmental chemicals (EPA 2014m). v-Embryo models are initially being developed to assess developmental effects in the eye, blood vasculature, genital tubercle, and limb. These systems have many canonical signaling pathways relevant to other organs and tissues. The models are developed based on developmental toxicology expertise and in-house assay data from ToxCast, ToxRefDB, stem cells, and zebrafish. v-Embryo models have already demonstrated their utility as hypothesis-testing platforms and for organizing the extant data within a systems biology framework. This framework is one that represents key events, accounts for interactions at different levels of

biological organization, and can address multiple kinds of stressors and exposure regimens. Model development is still in early stages, but the models are considered to be one of the more promising approaches to providing rapid and accurate health effects assessments in the future (DeWoskin et al. 2014; Knudsen and DeWoskin 2011; Knudsen et al. 2011b; Knudsen and Kleinstreuer 2011).

The goal of the v-Liver model is to construct a cell-based tissue simulator that uses systems models, a knowledgebase of chemical effect networks, toxicokinetic information, and *in vitro* data to predict chemically induced hepatotoxicity quantitatively from simulated exposures (EPA 2014n). The v-Liver model simulates hepatic functions by considering three main biological processes: (1) blood flow into the liver carrying nutrients and chemicals to cells, (2) molecular cross-talk networks that determine cellular responses, and (3) the dynamic interactions between cells that maintain homeostasis or result in histological effects (Shah and Wambaugh 2010). Blood flow is simulated by extending a PBPK model to calculate microdosimetry in the hepatic lobule using ordinary differential equations (Wambaugh and Shah 2010). Molecular cross-talk networks in individual cells are simulated using nondeterministic Boolean networks (Jack et al. 2011). Initially, the focus of the v-Liver model is to simulate key hepatocellular phenotypes in acute and chronic lesions such as hypertrophy, injury, death (necrosis/apoptosis/autophagy), Kupffer cell activation, or cell cycle progression. Many possible molecular events might lead to these cellular responses, and many of these events could be a consequence of nuclear receptor activation. Evidence from the literature is being organized on molecular and cellular perturbations by nuclear receptor activators, including xenobiotic and endogenous metabolism, oxidative stress, mitochondrial injury, DNA damage, the cell cycle, and apoptosis.

Extensive work supported in part by the Department of Defense has focused on building 10 different virtual models or "human organs-on-chips" and will provide an additional and potentially highly useful source of data for the VT models (Wyss Institute 2012). This effort is designed to streamline the drug development process and more effectively predict safety of drugs and chemicals in humans.

Virtual models are also briefly discussed in Section 4.4 as one of the new approaches that can address recurring issues in risk assessment, in this case, dose-response characterization.

3.3.2.5 HT Exposure Estimation: ExpoCast Prioritizations

The use of HT assays to characterize biological activity *in vitro* enables prioritization of potential environmental hazards once the results of *in vitro* assays have been anchored to, and found to be predictive of, *in vivo* effects. Without capabilities for HT assessment for potential exposures, prioritization (with respect to potential risk) cannot be completed, as most chemicals have little or no exposure data (Arnot and Mackay 2007; Arnot et al. 2010a; Arnot et al. 2010b; Cohen Hubal et al. 2010; Goldsmith et al. 2014; Hubal 2009; NRC 2006; Rosenbaum et al. 2008; Rotroff et al. 2010; Sheldon and Cohen Hubal 2009; Wetmore et al. 2012). Currently, few, if any, inexpensive *in vitro* assays are widely available to characterize the properties of chemicals that are relevant to exposure. Furthermore, studies assessing both the presence of environmental chemicals in the immediate vicinity of individuals (exposure potential), and any known biomarkers of actual exposure, are expensive, labor intensive, and, with the notable exception of CDC's NHANES,

typically difficult to extrapolate to the general population (Angerer et al. 2006; Eskenazi et al. 2003; Rudel et al. 2008). For these reasons, exposure prioritization must rely on mathematical models that when parameterized by chemical-specific properties, provide a structured, consistent way to approach large numbers of unknown chemicals.

Physicochemical properties (e.g., water solubility, preference for binding in lipids) inherent to a given compound have been used to predict potential bioaccumulation within ecological species to make HT prioritizations for potential chemical exposure (Gangwal et al. 2012; Reuschenbach et al. 2008; Walker and Carlsen 2002; Walker et al. 2002). Environmental fate and transport models are designed to account for the accumulation of compounds in various environmental media (i.e., air, soil, water) and for the degradation rates of those compounds in those media. These fate and transport models enable predictions of human exposure based on assumptions of human interaction with environmental media and derivation of food from the environment (Arnot and Mackay 2007; Arnot et al. 2010a; Arnot et al. 2010b; Rosenbaum et al. 2008). Parameterized based on chemical structure and production volumes alone, these models can be used to make HT exposure prioritizations (Arnot and Mackay 2007).

EPA initiated an ExpoCast program for exposure model development and prioritization. The framework is designed to be flexible and expandable to incorporate new HT exposure models as they become available. Two quantitative fate and transport models amenable to HT operation have been developed: USEtox (Rosenbaum et al. 2008) and RAIDAR (Arnot and Mackay 2007). These models have been empirically assessed for their ability to predict exposures inferred from the NHANES data set. More recently, Wambaugh et al. (2013) proposed a framework for HT exposure assessment, and demonstrated applications with an analysis that predicted human exposure potential for chemicals and estimated uncertainty in these predictions by comparison to biomonitoring data. The far-field mass balance human exposure models (USEtox and RAIDAR) were used in conjunction with an indicator for indoor or consumer use to evaluate 1936 chemicals. The model predictions were compared to exposures inferred by Bayesian analysis from urine concentrations for 82 chemicals reported in NHANES. Joint regression on all factors provided a calibrated consensus prediction, the variance of which serves as an empirical determination of uncertainty for prioritization on absolute exposure potential. Information on use was found to be most predictive; generally, chemicals above the limit of detection in NHANES had consumer/indoor use.

NexGen efforts to incorporate exposure prioritization information could proceed along three fronts. First, efforts to evaluate the utility of the predictions would be undertaken to determine if the chemicals of highest priority are indeed present in the environment. Next, further model development is needed for fate and transport predictions—specifically for exposure from personal contact sources (i.e., consumer use). Third, model results could be used to estimate mg/kg body weight/day (accompanied by uncertainty characterization) for application in calculating risk-based prioritizations.

3.3.2.6 HT Assays to Evaluate Thyroid Pathway Disrupting Chemicals –Workgroup Recommendations

EPA's NexGen Thyroid Disrupting Chemical Workgroup (EPA 2013a) conducted a thyroid prototype case study that reviewed existing ToxCast assays and provided recommendations for how the data could be used to predict thyroid disruption-induced developmental neurotoxicity. A major reason the workgroup selected the thyroid hormone system as its prototype is that the underlying biology of thyroid hormone homeostasis is well established, thus enabling the elucidation of the pathway(s) for thyroid hormone disruption (Zoeller and Crofton 2005). The workgroup identified three issues to address for HT assay use to predict chemically induced developmental neurotoxicity via disruption of thyroid hormone homeostasis: (1) assay identification and refinement; (2) algorithm development for toxicity and hazard prediction; and (3) standards development for assay conduct, data analysis, and data reporting for risk assessment needs. Following is a brief summary of the case study findings with respect to these issues.

Assay Identification and Refinement

As a first step, the workgroup identified the HT assays in the ToxCast database that assess endpoints known to be relevant to disruption of thyroid function. ToxCast contains multiple assays relevant to assessing the potential for a chemical to disrupt thyroid hormone homeostasis. Coverage of the effects of concern, however, is quite variable. Although five of the identified assays evaluate endpoints that directly affect the thyroid hormone pathway (e.g., thyroid hormone receptor binding and thyrotropin-releasing hormone receptor binding), the rest evaluate endpoints not specific to the thyroid hormone pathway. For example, of the 90 assays identified as thyroid relevant, 85 are related to hepatic stimulation, metabolism, and clearance of thyroid hormones. Alteration of these pathways influences thyroid hormone homeostasis indirectly. Neurodevelopmental effects via thyroid disruption by this mechanism are thus secondary effects of a chemical (e.g., inadequate hormone availability due to increased elimination). Secondary effects contrast with primary effects, whereby a chemical interferes directly with the function of the thyroid gland itself or interacts at the site of thyroid hormone receptor in the brain of a developing organism.

Adequately assessing the potential of an environmental chemical to disrupt thyroid hormone homeostasis requires that appropriate endpoints be identified and assays developed and incorporated into testing schemes. Work is ongoing to identify the specific endpoints in the pathways that need to be tested. Additional assays not currently part of ToxCast need to be developed. A recent workshop review by Murk et al. (2013) provides a state-of-the-science assessment of important MIEs for thyroid disruptors, current and potential new assays for these MIEs, and recommendations for research priorities.

Algorithm Development for Toxicity and Hazard Prediction

The workgroup's second recommendation was to develop algorithms or decision logic flows that assess the potential adversity of the outcome and the uncertainty in the available data. Assays evaluating endpoints directly affecting the thyroid-related brain changes might be weighted more heavily in algorithms than those measuring upstream hepatic enzyme induction. Algorithms should address the possibility of multiple chemicals interacting with the same key event and one

interacting with various MIEs. Biological plausibility is also an important issue that should be addressed during algorithm development.

Incorporating many sets of dose-response information into combinatorial analysis requires some simplification of assay results. Many current HT assay results are simplified via classification as either a positive or negative ("hit" or "no hit") or are assigned a summary statistic such as an IC_{50} (the concentration producing a 50 percent inhibition of response) or lowest effective dose. Obviously, binary decisions such as hit/no hit determinations depend on the criteria chosen to define a hit. These criteria could be derived from statistical significance, biological significance, or an arbitrary, nominal level of change. Depending on the data set, the basis for the classification criteria might be difficult to determine, and might not be consistent across assays. Similarly, summary statistics depend on the model used to generate them or on the specific value chosen (such as IC_{50} versus IC_{10}, the concentration producing a 10 percent inhibition of response). Relative potency ranks also might vary depending on the shape of the dose-response curve, such that within a given set of chemicals, Chemical A could have the lowest IC_{50}, while Chemical B could have the lowest IC_{10} value. Lack of such information will lead to greater uncertainty in data use. Thus, these criteria need to be explicitly stated and accessible.

Assay Conduct, Data Analysis, and Data Reporting for Risk Assessment Needs

Understanding the characteristics of individual HTS assays and data used to screen chemicals for disruption of thyroid hormones is critical. Individual assays might be used in predictive algorithms or test batteries for hazard identification and prioritization. They also might be used to provide supporting data for individual chemical risk assessments. Although the uses are potentially diverse, several common assay characteristics will be needed. Minimally, the data reporting should include sufficient information to document assay conduct and reliability, the rationale for selecting exposure levels, data analysis techniques, and underlying assumptions regarding assay analysis, conduct, or conclusions.

Three advantages of the ToxCast data sets are the availability of (1) dose-response information for all assays, (2) assay method details, and (3) source code for all computational models used in the data analyses. Reliable dose-response information, transparency for the methods used, and reproducibility of the results (i.e., availability of model code and assay conditions) are critical for these types of assays to be useful in risk assessment.

In summary, the thyroid pathway case study was complicated by the multitude of target sites at which the thyroid axis could be disrupted (Crofton and Zoeller 2005; Murk et al. 2013); the secondary, indirect nature of the insult produced; and the complexity of the endpoint of concern—neurodevelopment. This case study was successful in identifying the nodes in the thyroid toxicity pathway that need additional assay coverage, the algorithm development and assay conduct issues that need to be addressed, and the data reporting requirements for using HTS data in an assessment.

3.3.3 Risk Assessment Implications Across the Tier 1 Prototypes

The Tier 1 prototypes provide examples of new HT approaches to develop data for screening and prioritizing huge numbers of chemicals (i.e., tens of thousands) into categories for focused research, further testing, or further assessment. The main approaches are QSAR modeling, read across, and HTS and HCS assays. Methods being advanced to support or interpret HTS/HCS data for use in risk assessment include toxicokinetic models (that relate *in vitro* doses to chemical concentrations at *in vivo* target sites), VT models (to provide an experimental and theoretical framework for integrating and interpreting HTS/HCS assay data, as well as other data types), and HT exposure data (used in conjunction with the toxicity potential to screen and prioritize potential risk). The development of the Tier 1 prototypes and experience with the HT approaches led us to the following inferences:

- QSAR models can be used to screen and prioritize a large number of chemicals, and in some cases to estimate LOAELs; criteria, however, are needed to characterize the confidence in the QSAR values for predictive accuracy relative to authoritative, traditional toxicity values; critical issues include QSAR model transparency and updated data for the training set.
- QSAR models, read-across, and HTVMD models have the potential to address some of the limitations in the current state of the HTS and HCS assays, for example, evaluating potentially toxic metabolites, volatiles, and limited-solubility compounds such as metals.
- Large sets of environmentally relevant chemicals can be screened in a diverse battery of *in vitro* assays; predictive models of toxicity can be developed using these data.
- *In vitro* PK data can support reverse PK models capable of extrapolating dose levels from the *in vitro* assays to equivalent *in vivo* rodent doses enabling initial quantitative comparisons between *in vitro* toxicity data and *in vivo* rodent toxicity data. Initial estimates of the equivalent human doses are also possible.
- Limitations in the HTS/HCS data include the need to assay larger numbers of chemicals (ToxCast Phase II data and beyond); variability or confounding factors in *in vivo* conditions (e.g., species, tissue, lifestage, metabolism, complex interactions), some of which are difficult to evaluate in *in vitro* systems, lessen the utility of HT approaches in predicting disease.
- At present, characterizing the results of HT testing as indicative of alterations in biological processes is generally more reasonable than as predicting disease.
- Cellular and multiscale biologically based models (e.g., VT modes) are needed to interpret Tier 1 HT data (as well as Tier 2 data) and to simulate the complex dynamics of multilevel biological organization and interactions. These models aim to capture spatial and temporal dynamics in AOPs and how chemical (or nonchemical) stressors can perturb normally functioning network controls along those pathways and cause disease.
- HTS/HCS *in vitro* data might be used along with HT *in vivo* data to develop new types of reference values that support or supplement traditional values (based primarily on *in vivo* animal studies), but further advances in methods are needed to develop HT quantitative values for use in risk assessment.
- The quality of the databases that support evaluations and associations between HT assay data and disease outcomes is central to improving the confidence in HT data predictions and the use of these data to support higher tier assessment values.

4 Advanced Approaches to Recurring Issues in Risk Assessment

In addition to supporting more rapid and efficient chemical-specific assessments as discussed above, new data types and advanced approaches are contributing to our understanding of recurring, cross-cutting issues in risk assessment. These issues are often sources of controversy due to limited data or lack of methodology. The issues discussed in this section include (1) individual versus population-level effects; (2) variability in human response due to a variety of factors (e.g., genetic differences, early-life exposures, toxicokinetic differences); (3) exposures to mixtures and nonchemical stressors; (4) interspecies extrapolation; (5) characterization of responses at environmental exposure levels; and (6) implications of new methods for addressing recurring issues in risk assessment. Additional details are captured in a series of NexGen-related published articles on human variability (Zeise et al. 2012); early-life exposure and later-life disease risks (Boekelheide et al. 2012); and multifactorial interactions of the environment and genes (Bell, S. and Edwards 2014; Patel et al. 2012a; Patel et al. 2012b; Shen et al. 2011; Smith, M. T. et al. 2011; Zhuo et al. 2012). Further relevant discussions are in the National Research Council's (NRC) reports *Toxicity Testing in the 21st Century* (NRC 2007b), *Science and Decisions* (NRC 2009) and, mostly recently in a NexGen paper by Krewski et al. (2014). The National Academy of Sciences' principles for uncertainty and variability analysis, articulated in *Science and Decisions* (NRC 2009), and reiterated in Appendix C of this report, are particularly relevant to these new approaches for risk assessment.

The application of new risk assessment methodologies that are key to the framework for the next generation of risk science has been explored in the context of the NexGen case study prototypes; this analysis indicated that many innovative methodological aspects of the NexGen framework are already beginning to be adopted in practice (Krewski et al. 2014). Of interest here is how new data types and approaches can inform these challenging issues and advance our ability to protect human health and the environment.

4.1 Individual versus Population-level Effects

Important to understand is that for environmental risk assessment, evaluating risks to the individual is not the same as evaluating risks to a population. In particular, an exposure effect at the level of the individual is a change in the ***magnitude*** of some measure of a toxicological effect for a given exposure level. An exposure effect at the level of the population is a change in the ***incidence*** effects of any particular magnitude, that is, the number of new cases in the population for that magnitude of effect within a specified period divided by the size of the population initially at risk.[41] The magnitude of change should be defined as it relates to severity, so that a greater magnitude represents a more severe effect. For instance, a decrease in body weight of 20 percent is greater in magnitude (and is more severe) than a decrease of 10 percent, and a "moderate" liver lesion is greater in magnitude (and is more severe) than a "mild" liver lesion. Thus, for a monotonic dose-

[41]Best presented as a ratio, as defined here, rather than just the number of new cases.

response relationship in an individual for any given endpoint, a higher exposure will lead to effects that are greater in magnitude and, thus, greater in severity. In a human population, increasing exposure levels will result in more individuals (i.e., higher incidence) at or above a given magnitude (severity) of effect for the endpoint considered. Increased exposure also will result in a greater magnitude of effects for a fixed percentile of the population. Thus, as magnitude of effect and incidence related to a given endpoint increase at the same time, more and more subjects will suffer from more and more severe effects (i.e., of larger magnitude) as exposure increases. Figure 27 illustrates this concept.

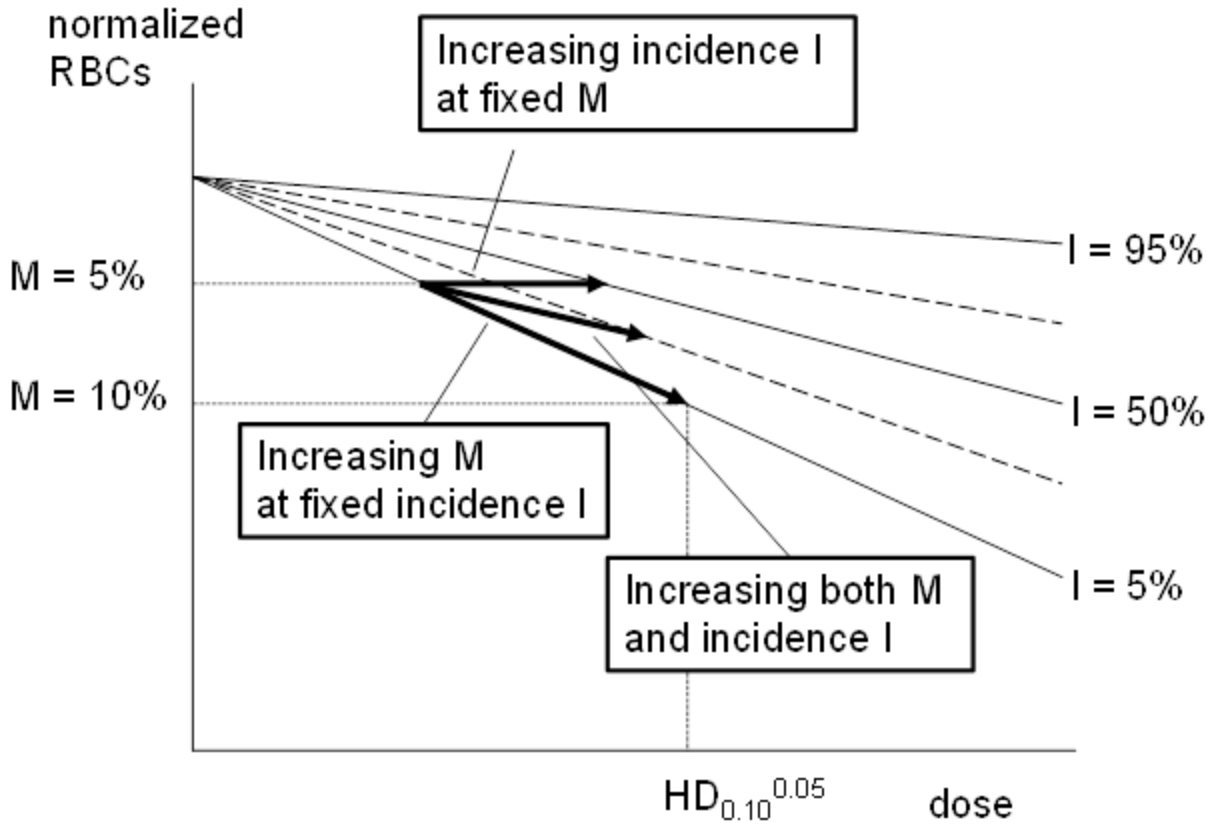


Figure 27. Magnitude (M) of the Effect and Incidence (I) for Decrease in Red Blood Cell Counts: Both Increase with Dose.

The solid middle line reflects the hypothetical dose-response relationship for decrease in red blood cells in the median individual (hence, I = 50%), the solid bottom line that of a more sensitive individual (at the 5th percentile of the population), and the top solid line that of a less sensitive individual (95th percentile). The dose-responses are normalized to each individual's background value on the y-axis. For a given effect size, for example, M = 5% decrease in red blood cells, a higher dose will result in a higher incidence (see shortest arrow). For a given percentile of the population, for example, I = 5%, a higher dose will be associated with a larger effect size M (see longest arrow). Similarly, a higher dose also can be associated with a simultaneous increase in I and M (see middle arrow). HD_{10}^{05} represents the human dose (HD) at which a 10% (or greater) magnitude (M) of effect is experienced at a population incidence (I) of 5%, a notation that is explained at the end of this subsection.

To evaluate uncertainties explicitly and quantitatively, the distinction between magnitude (or severity) and incidence should be maintained explicitly in a hazard (or risk) characterization. For example, when the aim is to derive a human limit value, the associated[42] target human dose is defined as a function of both the magnitude of the effect and the fraction of the population with that effect. For convenience, human dose or exposure is denoted HD, the magnitude of effect is denoted M, and incidence is denoted I. Their relationship is denoted as follows:

HD_M^I = human dose

where a fraction I of the population shows an effect of magnitude (or severity) M or more (for the critical endpoint considered).

This notation indicates the (estimated) human dose with the specified magnitude of effect and incidence, given the magnitude of effect. A major advance of this framework is the specification of HD_M^I as the final goal of hazard characterization because, in the past, the distinction between severity and incidence has usually not been made explicit. Specification of the value of M for different types of endpoints is discussed in the next two subsections.

4.2 Human Variability and Susceptibility

Human response to environmental chemicals is influenced by both intrinsic (e.g., genetics, lifestage, internal dosimetry) and extrinsic (e.g., chemical exposure, stress, nutrition) factors. New methods to examine gene-gene, gene-environment, and epigenome-gene-environment interactions are available (Baker 2010; Cordell 2009; Lvovs et al. 2012; Meissner 2012; Patel et al. 2012a; Patel et al. 2013; Patel et al. 2012b; Thomas, D. 2010). Zeise et al. (2012) explored how these factors can influence each biological and physiological step in the source-to-outcome continuum, and contribute to variability in the final health outcome (see Figure 28). The Zeise et al. (2012) review was informed by an NRC workshop, "Biological Factors that Underlie Individual Susceptibility to Environmental Stressors and Their Implications for Decision-Making." The authors considered both current and emerging data streams that are providing new types of information and models relevant for assessing interindividual variability.

In risk assessment, human variability typically is accounted for by including an uncertainty factor of 1, 3, or 10 in the calculation of a reference dose for noncancer health effects. Variability is not explicitly accounted for in cancer health assessment except for the incorporation of an age-specific adjustment factor of ≤10 for childhood exposures to genotoxic carcinogens. Rather, current cancer risk assessment approaches aim to account for sensitive subpopulations by using a 95 percent upper confidence limit in calculating estimates of potency. In a few cases, data on sensitive populations (e.g., asthmatics and those sensitive to air pollutants) might be specifically

[42]Note that a health-based guidance value derived in a hazard characterization would not be the same as the target human dose, but instead would be a (conservative) estimate of it.

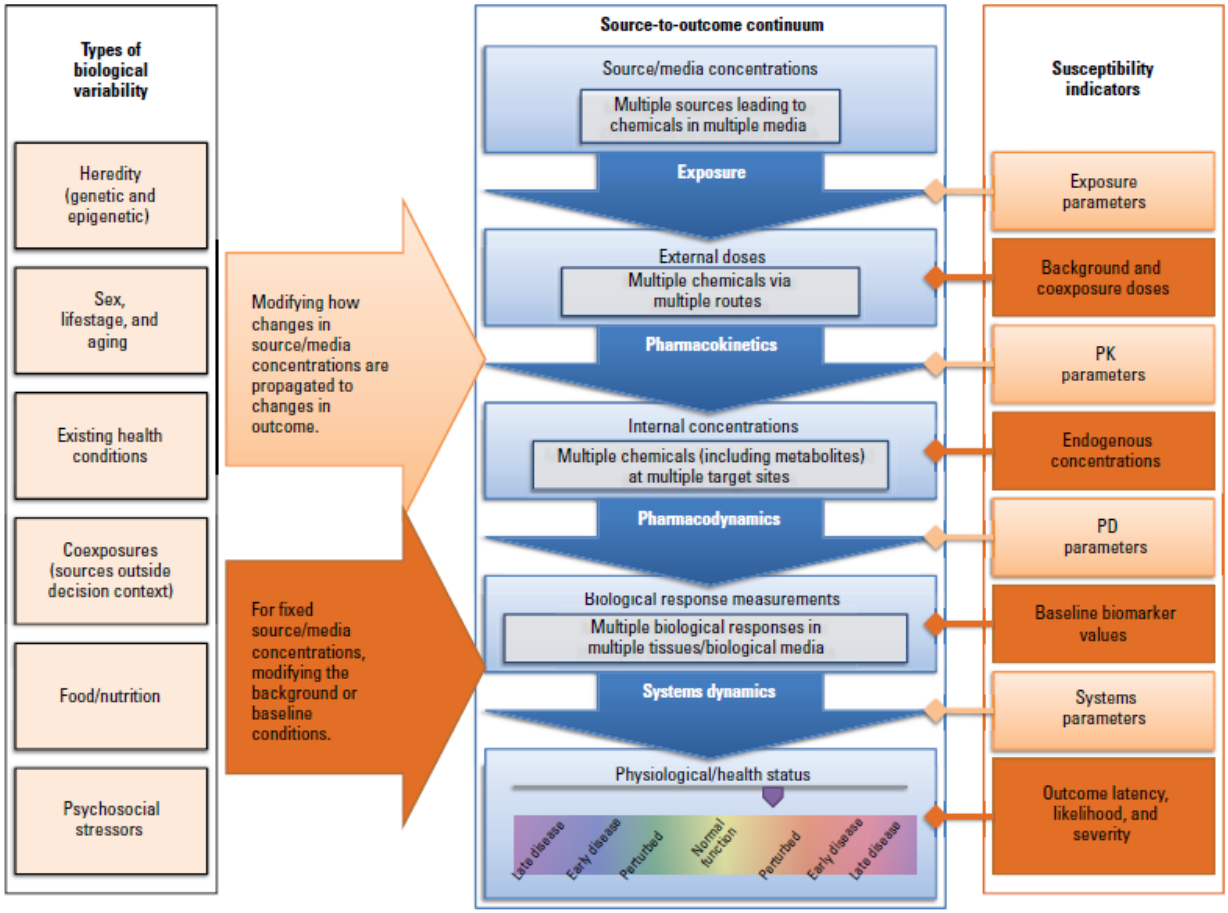


Figure 28. Framework Illustration of How Susceptibility Arises from Variability.
Multiple types of biological variability intersect with the source-to-outcome continuum, either by modifying how changes to source/media concentrations propagate through to health outcomes, or by modifying the baseline conditions along the continuum. The aggregate result of these modifications is variability in how a risk management decision affects individual health outcomes. The parameters and initial conditions along the source-to-outcome continuum serve as indicators of differential susceptibility, some of which are more or less influential to the overall outcome (see Figure 25 in original source) (Zeise et al. 2012). Reproduced with permission from *Environmental Health Perspectives*.

incorporated into risk assessments. Figure 29 from Zeise et al. (2012) illustrates how different types of variability can influence dose-response relationships.

The following discussion addresses factors that contribute to variability in human response to environmental exposures, and how new data and approaches will reduce uncertainty in estimating risks.

4.2.1Genomic Variability

Understanding the interaction between genetic and environmental factors will greatly improve our ability to estimate and manage public health risks. An estimated 20–50 percent of phenotypic variation is captured when all single nucleotide polymorphisms (SNPs) are considered simultaneously for several complex diseases and traits. The proportion of total variation explained by individual genome-wide-significant variants has reached 10–20 percent for several diseases (Visscher et al. 2012). Copy number variation and unexplored noncoding ribonucleic acids (RNAs),

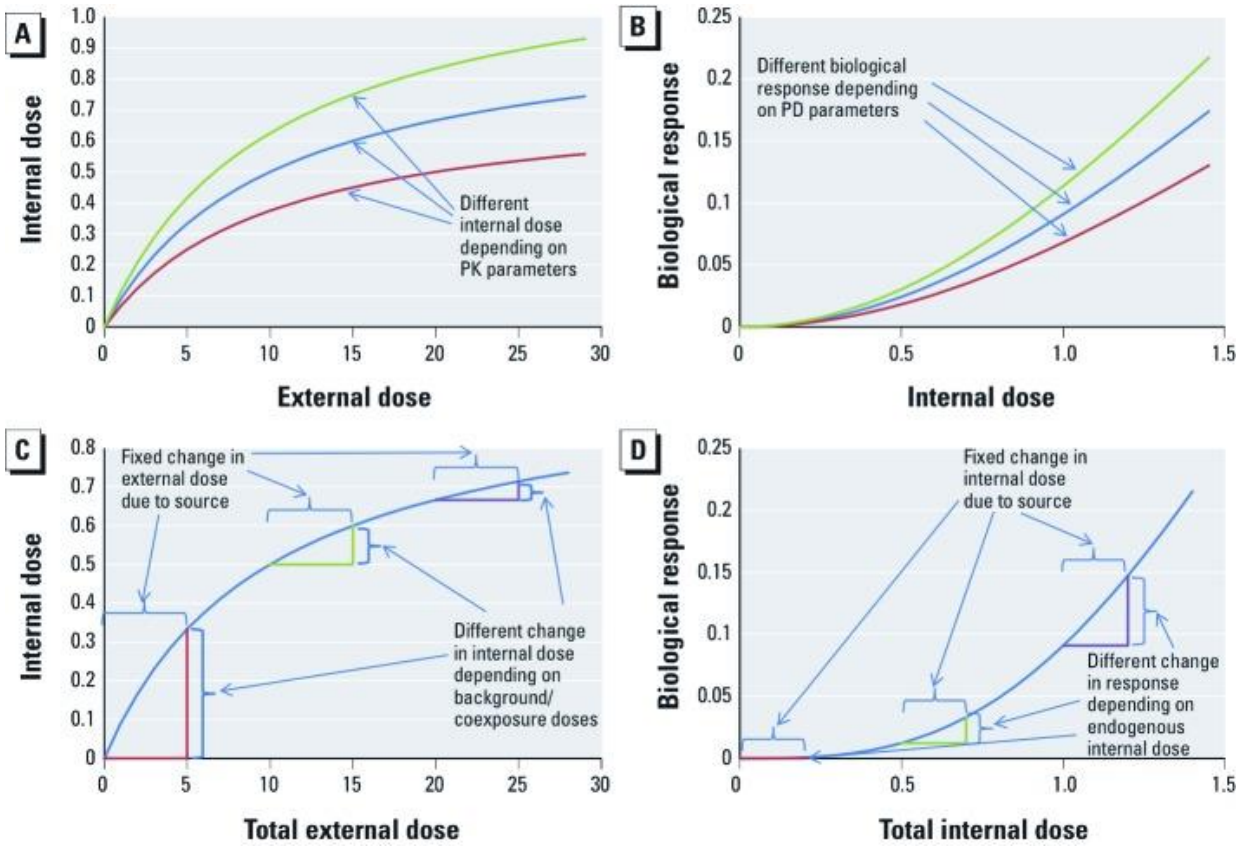


Figure 29. Effects of Variability in (A) Pharmacokinetics (PK), (B) Pharmacodynamics (PD), (C) Background/ Exposures, and (D) Endogenous Concentrations.
In (A) and (B), individuals differ in PK or PD parameters. In (C) and (D), individuals have different initial baseline conditions (e.g., exposure to sources outside of the risk management decisions context; endogenously produced compounds) (Zeise et al. 2012). Reproduced with permission from *Environmental Health Perspectives*.

microRNAs, and epigenetic factors most likely also contribute to human variability. Environmental factors are thought to contribute the remaining variability. The term "environmental factors" as used here is broadly defined to include diet, exercise, chemical exposures, and other factors.

Several approaches to generating and evaluating genomic data are emerging that can provide new insights into human variability. These include (1) computational modeling approaches in which variability in parameter values is simulated and differences among subpopulations is explored (Diaz Ochoa et al. 2012; Knudsen and DeWoskin 2011; Shah and Wambaugh 2010); (2) high-throughput *in vitro* data generation using cells lines with different genetic backgrounds (Abdo et al. 2012; Lock et al. 2012; O'Shea et al. 2011); (3) *in vivo* studies in genetically diverse strains of rodents to identify genetic determinants of susceptibility (Harrill et al. 2012; NIEHS 2014b); (4) comprehensive scanning of gene coding regions in panels of diverse individuals to examine the relationships among environmental exposures, interindividual sequence variation in human genes, and population disease risks (Mortensen and Euling 2013; NIEHS 2014d); (5) genome-wide association studies (GWAS) to uncover genomic loci that might contribute to human risk of disease (Abecasis et al. 2012; Bush and Moore 2012; NHGRI 2014a; Wright et al. 2012); and (6) association studies that correlate measures of phenotypic differences among diverse populations with

expression patterns for groupings of genes based on coexpression (Friend 2013; Patel et al. 2012a; Patel et al. 2013; Weiss et al. 2012). New understanding of the contribution of epigenomics to disease is rapidly advancing with evaluation of changes such as differential methylation of deoxyribonucleic acid (DNA) (Hansen et al. 2011; Rakyan et al. 2011; Teschendorff and Widschwendter 2012).

The approach reported by Lock et al. (2012) is being used in Tox21 Phase II (in collaboration with Rusyn and colleagues at the University of North Carolina) to expand the study of interindividual differential sensitivity to 180 toxicants. The researchers are evaluating approximately 1100 distinct human lymphoblastoid cell lines, with densely sequenced genomes representing 9 races of humans. Data will be collected on more chemicals in the future, and the numbers of chemicals evaluated in this manner will expand. The large number of human cell lines used allows for an analysis of determinants associated with differential cytotoxicity *in vitro*. Panel "a" in Figure 30 illustrates one example of how these new types of genetic variation data can be used in risk assessment, in this case, how a population concentration-response curve can be estimated for cycloheximide based on HT *in vitro* data using human cell lines with different genetic backgrounds. Although differences between immortalized cell lines and *in vivo* cells should be considered when interpreting results of this type, this approach can provide significant new insights into variability in human response and better inform current and future risk assessments. Other examples of human variability data are discussed in the benzene prototype (Section 3.1.1) and in Box 10 using GWAS data.[43]

The Tier 3 prototype for benzene-induced leukemia and the example presented in Box 10 illustrate how identifying gene networks and interactions advance our understanding of disease progression and the causal nature of gene/pathway alterations in leukemia. This knowledge will enhance our ability to screen chemicals having limited health effects data for their potential to increase risks of a specified disease if they are found to cause similar mechanistic disruptions. Risk assessments of the future will increasingly incorporate these types of data to replace uncertainty factors and to improve risk management for susceptible subpopulations.

[43]The differential risks conferred by human genetic variability are complex and might not be captured by analyzing small-scale gene variability alone. Hundreds to thousands of genes are likely to be involved in any disease, and multiple variations in genetic makeup might confer similar increased or decreased risk for the same disease. Disease occurrence also could be influenced by emergent system properties that require analysis of not only how gene variations affect cellular components, but how effects on critical network interactions propagate through higher levels of the biological system (Torkamani et al. 2008). Consequently, although incorporation of new types of data can help improve characterizations of human variability, the characterizations are likely to be incomplete.

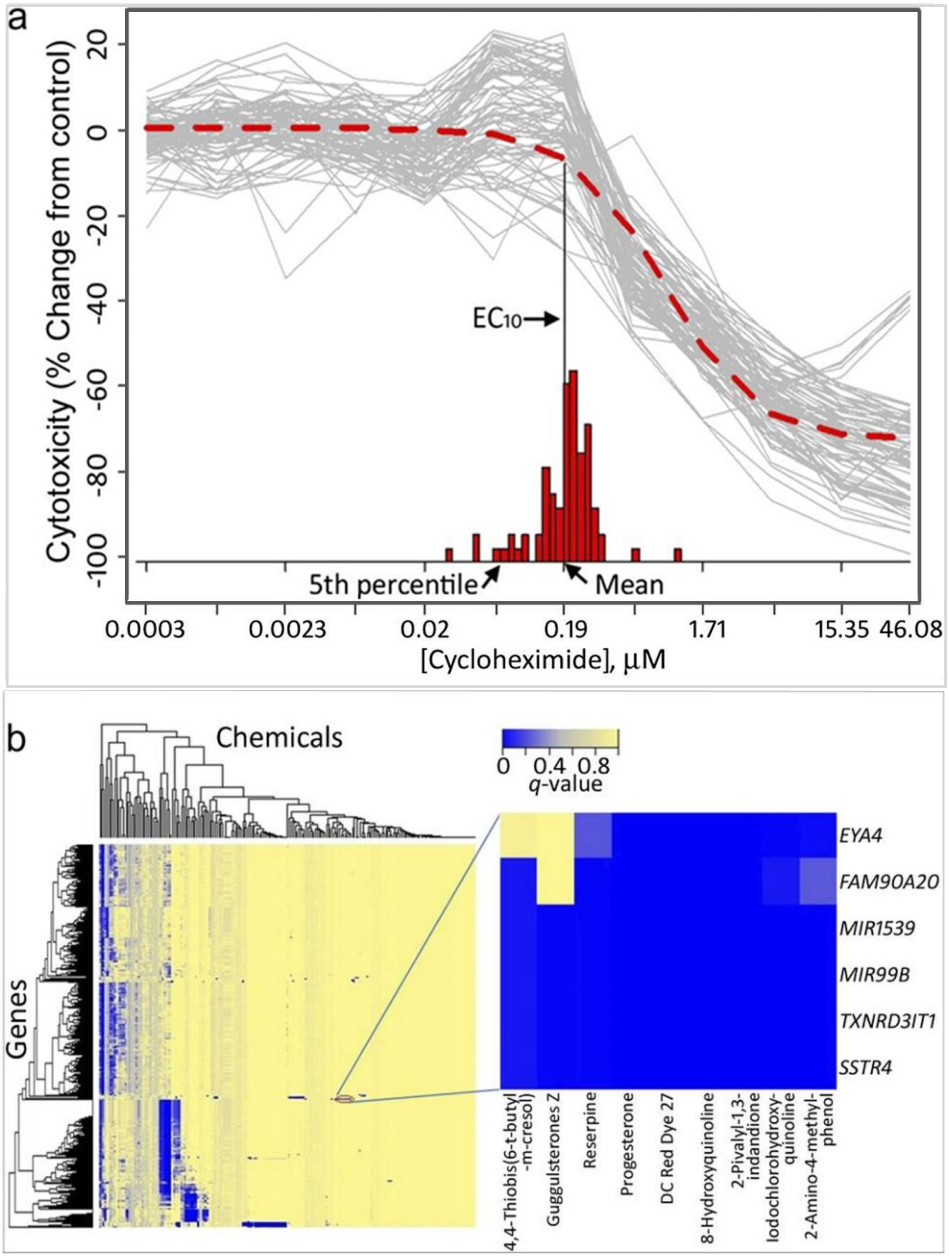


Figure 30. New Types of Genetic Variation Data Can be Used in Risk Assessment. Panel a: Population concentration-response was modeled using *in vitro* quantitative high-throughput screening (qHTS) data and cycloheximide data (cytotoxicity assay) as an example. Logistic dose-response modeling was performed for each individual to the values shown in gray, providing individual 10% effect concentration values (EC_{10}). The EC_{10} values obtained by performing the modeling on average assay values for each concentration (see frequency distribution) are shown in the inset. Panel b: A heat map of clustered false discover rates (q values, see color bar) for associations of the data from caspase-3/7 assay with publicly available RNA-Seq expression data on a subset of cell lines. A sample subcluster is shown (Lock et al. 2012). Reproduced with permission from *Oxford Journals*.

Box 10. Combining Genetics and Bioinformatics to Improve Estimates of Variability in Human Response

Variability in human response to chemical exposures is partly due to genetic influences. The National Center for Biotechnology Information at the National Library of Medicine in the National Institutes of Health has a vast array of databases devoted to human variability, especially genotype to phenotype associations. These resources include dbSNP (database of single nucleotide polymorphisms and estimates of their occurrence within the population); dbGaP (database of Genotypes and Phenotypes); GTEx database (Genotype Tissue Expression); OMIM (Online Mendelian Inheritance in Man); and PheGenI (Phenotype Genotype Integrator, which aggregates information from many of the aforementioned resources).

In this example, genome wide association study (GWAS) data were reviewed to examine the relationship between genotype and white blood cell count in benzene exposed and non benzene exposed workers in China. This work has been used, in part, to describe a possible mode of action for benzene hematotoxicity. Lan et al. (2004) identified single nucleotide polymorphisms (SNPs) associated with four DNA repair and genomic maintenance genes that could be involved in carcinogenesis. These SNPs confer significant odds ratios from 1.4 to 5.7 of having a white blood cell count <4000 cells/µL blood. This observation demonstrates a quantitative increased risk of hematotoxicity in individuals with any of these SNPs. Hematotoxicity is highly correlated with leukemia resulting from benzene exposure. Hence, these SNPs also might confer susceptibility to leukemia.

PheGenI provides links to dbSNP to view genetic diversity of SNPs within reported populations. For instance, rs12951053 s A/C genotype is reported to occur in 51.1% of the Chinese population and 31.1% of the Japanese population; and among Europeans and those of European descent, the A/C genotype occurs in approximately 9 17 % of the population (NCBI 2014c).

Overall, the minor allele (C), has a relatively low penetration within the global population at just 18.7% ± 2.2% (mean ± standard error of the mean), and an average heterozygosity of 30.0% ± 24.5% (average ± standard error of the mean).

Using the global minor allele rate of 18.7% ± 2.2 %, a probability function and model can be constructed so that any given member of the population has the minor allele A for rs12951053 SNP. Using this probability function, the number of people who might have a white blood cell count <4000 cells/µL blood can be estimated, and thus the potential for hematotoxicity. Model uncertainty also can be estimated. This approach thus provides a quantitative estimate of human health hazard.

In addition, this approach can help inform the analysis of environmental justice issues. For instance, by using census demographic data and the SNP occurrence data for people of particular races or other specific groups, creating probabilistic models that might more accurately reflect the SNP pool of a population, and thus, human variability, is possible. For at risk populations, regulatory agencies could use this type of information to inform their site specific risk assessments, such as a Superfund site risk assessment in the United States.

4.2.2 Early-life Exposures

Early-life exposures to chemicals can invoke molecular effects that appear to result in increased susceptibility to disease or other morbidity later in life, often via epigenetic modifications (Boekelheide et al. 2012). Evidence from both humans and animals helped establish the influence of early-life exposure on later-life outcomes. For example, human observational data and animal studies report that arsenic exposure during prenatal and early postnatal life increases the risk of cancer, respiratory and cardiovascular diseases, and neurobehavioral disorders (Boekelheide et al. 2012; Cronican et al. 2013; NRC 2011; Tokar et al. 2012; Tokar et al. 2011). Later-life outcomes can be influenced by time of exposure, predisposition of a species to a particular disease, an individual's genetic predilection to disease, or gender. Improved ability to predict disease risk associated with *in utero* or early postnatal exposures results from advances in identifying the targeted genomic region of chemicals and chemical mixtures, epigenetic alteration of gene expression, and the causal links between early-life chemical exposure and later-life outcomes (Boekelheide et al. 2012; NRC 2011).

Computational and statistical models for developmental effects provide valuable new approaches predicting risks from *in utero* exposures. The Tier 1 sections present examples of developmental

toxicity models based on the ToxCast data (Kleinstreuer et al. 2013a; Kleinstreuer et al. 2011a; Knudsen et al. 2009; Sipes et al. 2011a) and discuss the systems biology models under development in the v-Embryo project (DeWoskin et al. 2014; Knudsen and DeWoskin 2011; Knudsen et al. 2011b; Knudsen and Kleinstreuer 2011). These new approaches and supporting data are advancing our ability to understand normal developmental biology, and to predict how chemical perturbations can lead to adverse outcomes, especially when addressing the very challenging issues around assessing effects from *in utero* exposures (e.g., rapidly changing kinetic and dynamic processes in the developing infant, critical windows of exposure, and sparse data).

Epigenetic biomarkers for early-life exposures (e.g., placental epigenetic biomarkers, plasma biomarkers) could be used as early indicators of adverse health effects later in life. Development and interpretation of epigenomic[44] biomarkers are in early stages (Hansen et al. 2011; Rakyan et al. 2011). As our understanding of the underlying epigenetic mechanisms advances (e.g., DNA methylation, histone modification, microRNA), however, our ability to use biomarkers of early-life exposure to predict later-life disease risk will improve. A good example is the work based on associations between early-life exposure to arsenic and DNA hypomethylation, with the subsequent development of arsenic-induced skin lesions (Boekelheide et al. 2012; Pilsner et al. 2009).

4.2.3 Variability in Internal Dosimetry

Differences in individual absorption, distribution, metabolism, and excretion rates (i.e., toxicokinetics [TK][45]) for any given chemical will affect the levels of the chemical found in different parts of the body, including at its proposed target site, the main value of interest in hazard assessment. The uncertainty factor mentioned in Section 4.1 is used to calculate a reference dose for noncancer health effects to account for human variability and has two parts—one for pharmacodynamic (PD) differences and one for pharmacokinetic (PK) differences. The PK portion of the uncertainty factor for interindividual variability is 3.16 ($10^{1/2}$). When PK data are available, physiologically based pharmacokinetic (PBPK) model results are used to estimate the internal dosimetry of chemicals for any given exposure and route, and replace the uncertainty factor.

Extensive literature is available on the general use of PBPK models in risk assessment (Clewell et al. 2002; EPA 2006; McLanahan et al. 2012; WHO 2010) and, more specifically, use of models along with advanced statistical approaches to characterize population variability (Barton et al. 2007; Chiu et al. 2009).[46] A recent analysis of population distributions for PK parameters affecting chemical

---

[44]The "omic" in epigenomic is in reference to data on a complete range of epigenetic biomarkers (i.e., the whole picture). **Epigenetic** refers to the kind of change in gene activity that the marker represents.
[45]**Toxicokinetics** (TK) – Risk assessors will sometimes use the word "toxicokinetics" (TK) to distinguish the chemical as a toxicant from a drug and the more traditional use of the word **pharmacokinetics** (PK). The root word ''pharmakon'' has complex meaning that encompasses both a remedy and a toxicant (and more broadly any biologically active substance). Both terms are in common use, and appear in the text. They relate to the same processes, and are interchangeable.
[46]Bois and Clewell (2010) provide a particularly good presentation of the determinants of population heterogeneity and the intercorrelation of covariates affecting a chemical's clearance from the body.

disposition supports the use of the default value of 3.16 to account for interindividual variability when the toxin of interest is the parent compound, and for the most sensitive subpopulations, except for very young children (younger than approximately 3 months) (Valcke and Krishnan 2014). When the probable toxin is a metabolite, however, or when risk assessors have additional PK data (especially for susceptible subpopulations) that can be incorporated into a PBPK model, the model results provide a better characterization of the chemical's toxicokinetics and are used in lieu of the uncertainty factor to reduce uncertainty in the derived reference value.

In vitro-to-*in vivo* extrapolation (IVIVE) and reverse dosimetry (RTK) are central to the use of the new HT data in NexGen assessments. As highlighted in the Tier 1 assessment discussion, IVIVE and reverse dosimetry are being used to estimate *in vivo* exposures and internal concentrations (Rotroff et al. 2010; Wetmore et al. 2012). This information is essential to apply the HT results for relevance in humans, specifically for the information needed to characterize dose-response (i.e., external dose [estimated from RTK], internal concentration [estimated from IVIVE], and the associated effects [from the HT assay and systems modeling results]). As with PBPK models, the main limitation in the application of IVIVE and RTK approaches is the availability of data to support the critical PK parameter values for rates (and sites) of absorption, metabolism, elimination, and tissue partitioning.

The data developed for PBPK models will complement and extend the domain of applicability for IVIVE and RTK models, and vice versa. Concerted efforts to populate databases for needed PBPK model parameter values historically have employed various structure-activity relationship (SAR)/quantitative structure-activity relationship (QSAR) algorithms (Béliveau et al. 2003; Peyret and Krishnan 2011; Poulin and Haddad 2013), extrapolations from *in vitro* data (Harwood et al. 2013; Poulin and Haddad 2013), the more resource intensive compilations and curations of literature (DeWoskin and Thompson 2008; Hines 2007, 2013; Thompson et al. 2009), or targeted *in vivo* studies. These data resources can be used to assist the IVIVE and RTK effort. Conversely, the focused interest in developing IVIVE and RTK parameter values on the much larger domain of chemicals than traditionally addressed with PBPK models is likely to add a significant amount of new data and methodology that will benefit PBPK modeling and PK approaches in general. As with many such efforts, advances in the modeling depend on the free exchange and availability of these data resources.

4.3 Mixtures and Nonchemical Stressors

Cumulative risk addresses exposure to combined threats from all intrinsic and extrinsic stressors (e.g., chemical exposure, pharmaceutical use, underlying susceptibility, socioeconomic status, work-life stress) and factors that improve health (e.g., good diet, exercise). Assessing cumulative risk remains a challenging area for human health risk assessment. Only a few studies have examined the potential impact of exposure to environmental chemical mixtures, or to mixtures and nonchemical stressors; while innumerable combinations of chemical mixtures and nonchemical stressors occur in the environment. Conventional methods for risk assessment have progressed little in overcoming this particularly daunting challenge. New methodologies in systems biology, computational models, and data mining are promising based on a more comprehensive disease-oriented approach to identifying and managing cumulative risk for chemical classes or structures.

Understanding and modeling common patterns of significant pathway or network alterations associated with disease are integral to developing efficient approaches for assessing risk from mixtures, specifically in evaluating how components in that mixture might alter specific nodes, and whether additive, antagonistic, or synergistic outcomes would be expected. The HTS and omics data support bioinformatic and computational efforts to characterize mixtures. HTS and omics assay data can be combined with bioinformatics data mining and computational cellular signaling simulations to predict possible disease outcomes (initially for screening-level assessments).

As our understanding continues to evolve on how nonchemical stressors affect network interactions and modulate disease, we will begin to address the very challenging assessment of potential cumulative chemical and nonchemical stressor impacts on health. Mixtures assessment logically would focus on anthropogenic and natural chemicals known to co-occur in the environment. Biomarkers of exposure data will play a key role in determining internal levels resulting from actual environmental exposures. Because epigenomic networks are more easily modulated by environmental factors than the genome, epigenomics will be the initial focus so that mechanisms that mediate cumulative risks imposed by exposures to environmental factors can be identified (Bollati and Baccarelli 2010; Cortessis et al. 2012; Koturbash et al. 2011).

4.4 Interspecies Extrapolation

A better understanding of toxicological or biological pathways and their similarity (or lack thereof) among species will improve our ability to extrapolate chemical effects across species and to select model organisms for testing (Aldenberg and Rorije 2013; Kenyon 2012; Lalone et al. 2013; NRC 2005; Smirnova et al. 2014).[47] Animal models in hazard identification and characterization of dose-response traditionally use chemical testing of mammalian species, and apply an interspecies (animal-to-human) uncertainty factor (≤10) or body-weight conversion factor to derive an EPA reference value. As knowledge increases on the extent of pathway conservation among species, alternative test species, including nonmammalian vertebrates (adult and embryonic zebrafish) and invertebrate models, will be more useful in chemical risk assessment. Regulatory toxicology as a whole will move toward increasing reliance on predictive approaches to assessing chemical risk, with greater emphasis placed on understanding chemical perturbation(s) of conserved biological pathways at key junctures, including molecular initiation events (MIEs) (e.g., activation or inactivation of specific receptors, enzymes, or transport proteins). As discussed in Section 3 (Box 3), an extensive effort to develop and interpret adverse outcome pathway (AOP) networks in terms of animal-to-human extrapolation is ongoing at the Organization for Economic Cooperation and Development (OECD 2014a).

Data from alternative mammalian species and *in vitro* models are valuable for both ecological and human health risk assessment when used in a pathway-based framework (Ankley et al. 2010). Extrapolation between species can occur at different levels of biological organization, such as the

[47]Pharmacokinetics is an equally important area for consideration. As cross-species extrapolation of pharmacokinetics has been discussed extensively elsewhere and is only mentioned here.

MIE, the pathway, and the organ or individual levels. Based on the similarity of pathway-based values to standard toxicological values, this approach appears to be useful for extrapolating hazard values across species, especially if a known pathway is involved.

That gene sequences are conserved—even between distantly related species—is well known, and conservation across species is indicative of an essential function. DNA sequence similarity can, but does not always, reflect a functionally conserved role for the genes in question. Investigations of gene function homology can be approached through interspecies comparisons of various components that affect the phenotype in question. The implicated genes, their sequence variation, and the relevant signaling pathways and tissues (cells, organs, circuits) are all informative. Thus, new approaches to understanding the underlying molecular mechanism can improve cross-species extrapolation (e.g., see Ankley and Gray 2013; Burgess-Herbert and Euling 2013; Chen, J. et al. 2007; Chiu et al. 2013; Jubeaux et al. 2012; Reaume and Sokolowski 2011).

4.5 Responses at Environmental Exposure Levels

New data and approaches are needed to resolve long-standing controversies about characterizing low exposure-dose-response relationships. Much discussion has been held in the risk assessment field about linearity versus nonlinearity, threshold versus nonthreshold responses, and cancer versus noncancer outcomes. With a few exceptions, available traditional studies have insufficient statistical power to inform responses at environmental concentrations, and more information is needed about AOP networks, variability in response, background levels (White et al. 2009), and dose-response model uncertainty (Slob et al. in press).

Risk assessors generally have relied on a combination of precedents and theoretical arguments with some mechanistic underpinnings to guide extrapolation approaches to low exposure levels. Although the tendency has been to compartmentalize cancer into linear, nonthreshold responses, and noncancer effects into nonlinear, threshold responses, the same mechanistic arguments below can apply to both:

- Clearance pathways, cellular defenses, and repair processes are thought to minimize damage so that disease does not result.
- Backgrounds of exposure or preexisting disease can result in additivity to preexisting response backgrounds.
- Statistically greater response variability in the human population (as compared to traditional inbred animal studies) flattens (i.e., linearizes) the low dose-response relationships (Crump et al. 2010b; Lutz 1990; NRC 2009).

Harmonization of the methods used to assess cancer and noncancer risk is critically important (Gaylor et al. 1999; NRC 2009). Many important biological pathways do not parse neatly into cancer or noncancer processes, rather disrupted biology can contribute substantially to both types of adverse outcomes. A holistic perspective (i.e., a systems approach) that accounts for progression of effects—and different spectra of effects as dose increases—is needed to incorporate and interpret the large amount of mechanistic information being generated by the health effects and medical research communities. These new data and this knowledge will help inform the low-dose range

issues in two primary areas: (1) a dramatic increase in data from laboratory and field (epidemiological) studies for response at low doses, and (2) the elucidation of mechanisms for response at low dose and dose progression. Of note is that many of the HTS assays in the ToxCast program use human cell lines, and a broad range of doses (some of which can be at levels comparable to expected environmental exposures) provides much more information on dose range-responses (Judson et al. 2014).

New experimental data to characterize dose-response relationships at environmental exposure levels will avoid extrapolations of higher doses that often are based on assumptions about the shape of the dose-response curve at low doses, rather than direct estimation of risk in the low dose region. New high-throughput experiments have resulted in a dramatic increase in the availability of dose-response data for many chemicals at environmentally relevant concentrations. Dose-dependent molecular changes associated with adverse outcomes now can be measured for hundreds (*in vivo*) to tens of thousands (*in vitro*) of chemicals (Judson et al. 2014; Rotroff et al. 2010; Sturla et al. 2014; Thomas, R. S. et al. 2011; Thomas, R. S. et al. 2012c; Thomas, R. S. et al. 2013c; Tice et al. 2013; Wetmore et al. 2013; Wetmore et al. 2012). Faster and less costly molecular epidemiology and clinical studies also provide valuable data on biological responses in environmentally exposed humans (McCullough et al. in press; Thomas, R. et al. 2014; Vineis et al. 2013). The power of an assay to detect an effect (assay sensitivity and experimental variability) will be an important determinant for the reliability of these direct empirical measurements.

Observed molecular changes include alterations in both magnitude and character of responses, reflecting underlying alterations in biology with increasing dose and time. Biological processes linked to disease that are consistently observed across the exposure range of interest are likely to be useful as biomarkers of exposure and effect (Institute of Medicine 2010; Thomas, R. et al. 2014). Observed molecular changes must be understood in a mechanistic context and in light of their impact on variability in human responses in the population.

Rhomberg et al. (2011) identified the challenge of translating modest degrees of underlying variation in biological response to discrete differences between healthy and diseased states. Specific molecular alterations have been shown to be causally related to (or be a risk factor for) a disease or multiple diseases, but more commonly individual changes act in concert to execute normal biology, adapt to insults, or lead to disorder and disease (Medzhitov 2008). Ultimately, knowledge of endogenous levels of a toxicant under study, background levels of other stressors, background incidence of disease, relevant biological/physiological pathways, and biological mechanisms for coping with toxicant stressors are all factors that must be taken into account in evaluating population dose-response. Although elucidating which dynamic changes are relevant to risks is challenging, incremental progress is being made.

NRC (2007b) recommended developing new approaches and models to generate the data needed for characterizing the dose-response curves and improving quantitative estimates of risk, especially at doses applicable to likely human exposures. Examples of some new approaches to dose-response modeling are described in Burgoon and Zacharewski (2008), Parham et al. (2009), Zhang et al. (in press), and Zhang et al. (2010b). The application of HT assays of pathway perturbations that directly measure biological effects at environmental exposure levels are described in Rotroff et al.

(2010) and Wetmore et al. (2012). The reduced cost of *in vitro* HT assays relative to *in vivo* toxicity tests enables the use of a much broader range of exposure levels, leading to a more detailed description of dose-response relationships throughout the exposure range of interest. Figure 31 summarizes the automated dose-response modeling approach proposed by Burgoon and Zacharewski (2008) and suggests how dose-response models could be developed using large-scale molecular biology studies.

Empirical dose-response models are used widely in health risk assessment. They will continue to be used in the near term for screening and categorizing toxic substances, determining toxic potency, determining a point of departure (POD) for low-dose extrapolation, determining human exposure guidelines, estimating risk under specific exposure circumstances, and interpreting human data.[48] Models that are based on a robust understanding of biological processes, in contrast, are less common, but are anticipated to become more so in the future. To date, the main biologically based models used in risk assessment are PBPK models (see Section 4.2). Well-developed and adequately tested PBPK models are currently used in risk assessment to simulate the toxicokinetics of a chemical or chemicals across dosing regimens (duration, amounts, delivery rate, routes) and species, or extrapolating from *in vitro* regimens to *in vivo* doses (IVIVE).

[48]Establishing human exposure guidelines for environmental agents involves determining a POD on the dose-response curve. Examples include a particular response level on a BMD model estimate of the dose-response, corresponding to a specified increase in risk usually in the 5–10% range, or a signal-to-noise-crossover dose introduced by Sand et al. (2011). This POD is reduced further by adjustment factors to derive a level of exposure considered to be protective of human health and the environment. NRC (2009) suggests an integrated approach to the establishment of human exposure guidelines using adjustment factors applied to the POD, where the magnitude of the factor depends on the "expected" behavior of the exposure-response curve at low levels of exposure. NRC also examined the influence of background exposures and background disease rates on the shape of the exposure-response curve at low levels of exposure.
Characterizing the expected response at low exposure levels (i.e., those the public is most likely to encounter) is another great challenge to previous methods used in risk assessment, specifically the use of relatively high-dose *in vivo* animal assays as the source of data for adverse health effects because the spectrum of adverse effects might be quite different at lower doses.

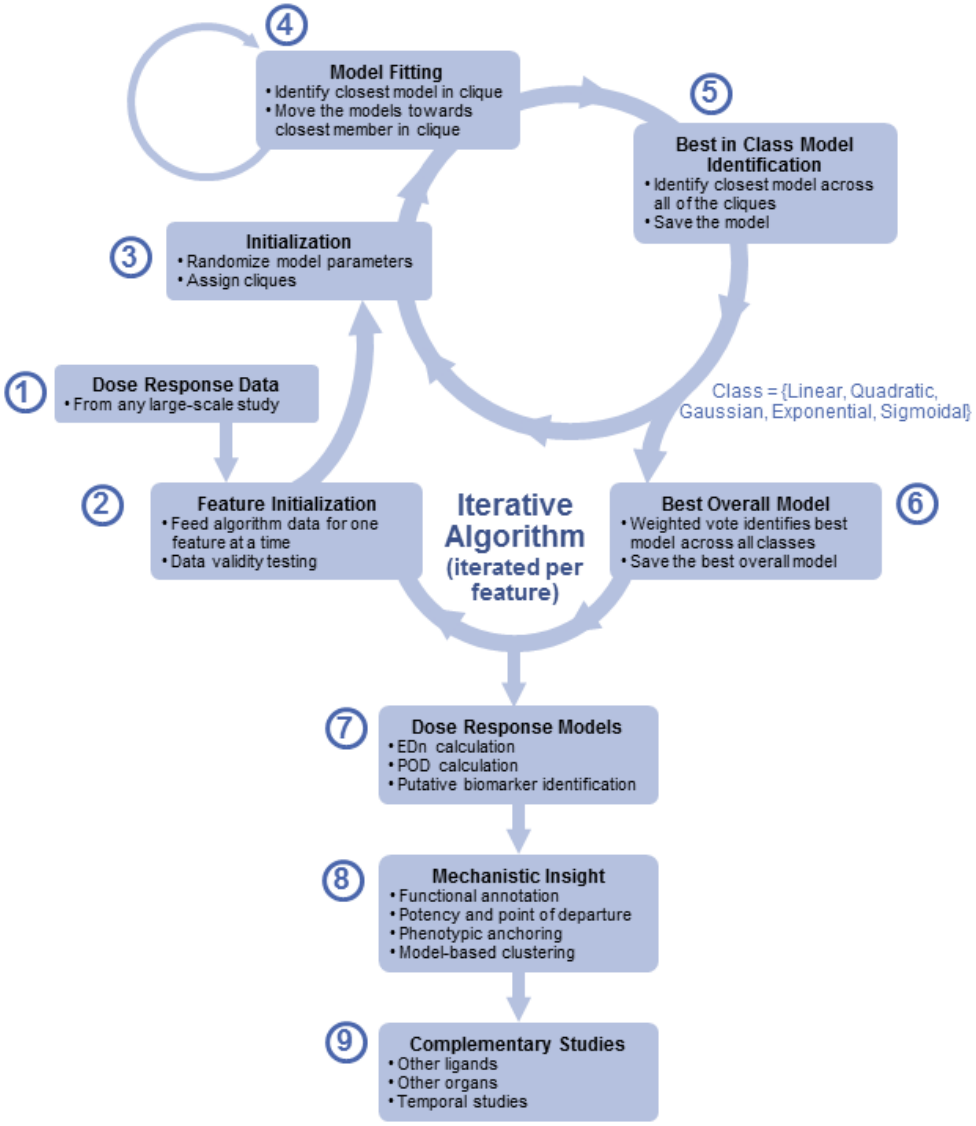


Figure 31. Overview of Automated Dose-response Modeling from Burgoon and Zacharewski (2008).
Step 1: Dose-response data from a large-scale study are loaded. Step 2: The application feeds dose-response data for one feature into the algorithm. Examples of feature data include messenger ribonucleic acid (mRNA), protein, or metabolite levels and enzyme or binding activities at each dose within a study. Step 3: The application initializes the particle swarm optimization (PSO) algorithm by randomizing model parameters and assigning cliques. Step 4: The PSO identifies the closest model in each clique at the end of an iteration and moves the members of each clique toward that model. Step 5: This iterative process ends once a best-fit model has been identified or when all iterations have been used. Steps 3 through 5 are repeated for each model class for the same feature, thus generating best-fit models for the linear, quadratic, Gaussian, exponential, and sigmoidal classes. Step 6: The best linear, quadratic, Gaussian, exponential, and sigmoidal models are compared with the best overall model using a weighted vote method. The model with the smallest Euclidean distance compared with the dose-response data receives the most votes. Step 7: The application uses the best overall model to calculate EDn and point-of-departure (POD) values, used to rank and prioritize putative biomarkers or chemical activities. Step 8: Model-based clusters can provide additional mechanistic insight by integrating potency and POD data with functional annotation and phenotypic anchoring. For example, EDn and POD data might generate model-based clusters for lipid metabolism and transport gene expression that could be associated with the occurrence of hepatic vacuolization and lipid accumulation. Step 9: Through complementary comparative studies using toxic and nontoxic congeners in responsive and nonresponsive species across time, data could emerge that differentiate biomarkers of exposure from toxicity-related responses that can support mechanistically based quantitative risk assessments. Reproduced with permission from *Oxford Journals*.

A new class of biologically based models called "virtual models" is being developed to simulate normal biology and to predict how chemical perturbations might lead to adverse effects (i.e., to predict a chemical's toxicodynamics) based on knowledge of potential mechanisms. Such models could be used to estimate the dose-response characteristics of a chemical for specific endpoints. Examples of virtual models being developed at various levels of biological organization or function include the:

1. European Virtual Physiological Human project (Hunter et al. 2010);
2. HumMod, a whole-body integrated human physiology model (Hester et al. 2011);
3. Virtual Cell (V-Cell), a spatially realistic quantitative model of intracellular dynamics (Moraru et al. 2008);
4. EPA's Virtual Embryo™ (v-Embryo) project, a suite of models that simulate normal development leading to the formation of blood vessels, limb-buds, reproductive systems, and eye and neural differentiation (Knudsen and DeWoskin 2011; Knudsen et al. 2011b);
5. EPA's Virtual Liver™ (v-Liver) model that simulates the dynamic interactions in the liver used to translate *in vitro* endpoints into predictions of low-dose chronic *in vivo* effects in humans (Shah and Wambaugh 2010);
6. Virtual Liver Network (German Federal Ministry for Education and Research 2014), a German initiative to develop a dynamic model of human liver physiology, morphology, and function integrating quantitative data from all levels of organization (Holzhutter et al. 2012); and
7. Hamner Institutes for Health Sciences' DILIsym® project that intends to identify new molecules that might cause liver toxicity and to understand the mechanism of existing toxicants (The Hamner Institutes For Health Sciences 2014).

In addition, the Physiome Project (Physiome Project 2014) is a major resource and model repository for hundreds of physiology models (Hunter et al. 2002).

Once fully developed, these models could dramatically improve our characterization of the dose-response relationship of various chemicals for several target tissues and functions.

4.6 Implications of New Methods for Recurring Issues in Risk Assessment

Based on the discussion above, and the examples provided throughout the report, the following summary inferences can be drawn about the use of new data and approaches in addressing recurring issues in risk assessment:

- Genetically derived **human variability and susceptibility** or resistance to environmental stressors can be evaluated using experimental *in vitro* and computational approaches; and emerging data streams (such as genetically defined human cell lines, genetically diverse rodent models, human omic profiling, and GWAS).

- New understanding of mechanistic events allows for greater **confidence in causal linkages** among exposure, molecular events, and adverse health outcomes; and enables the **identification and grouping of chemical mixtures and other environmental stressors** that can alter risk of a specific disease based on similarities of pathway perturbations.
- Omic events are well characterized across several species and **thus inform cross-species extrapolations**. Functional and omic responses that are highly conserved across many species facilitate cross-species considerations.
- New data types, collected in the range of environmental exposures, and systems models provide **better insights into low dose-response relationships** than previously possible. Mechanistic information on adaptive, maladaptive, and background responses will help characterize the shape of dose-response relationships for individuals and populations.

Based on the above, risk assessment likely will move to a more probabilistic description of risks derived from distributions of response across the human population, depending on several factors. Such factors include genetic makeup, lifestage, internal dosimetry, exposure to mixtures and other environmental stressors, and a better understanding of low dose-response relationships.

5 Lessons Learned from Developing the Prototypes

Perhaps the most critical revelation from the prototypes presented in this report is that the National Research Council (NRC) vision embodied in their report on *Toxicity Testing in the 21st Century* (NRC 2007b) can be realized, as evidenced by the remarkable progress in biology. Clearly, more must be done. Yet in the relatively few years since publication of that report, the focus of risk assessments has begun to shift from the traditional approach of using animal study data and uncertainty factors to the new assessment approaches demonstrated in the prototypes. The new approaches consider a different and broader array of data, a mechanistic understanding of adversity, and a move toward replacing uncertainty factors and extrapolations with data-derived probability distributions. This report provides additional scientific support for modernizing risk assessment.

Additionally, the methods discussed in the prototypes illustrate a convergence of perspectives and synergy of methodology occurring between the medical research community, traditionally focused on addressing treatment for clinically observable disease, and the toxicology community, focused on predicting outcomes from initial exposures. This convergence will greatly facilitate progress. Both communities are now developing and using tools and approaches to resolve the more detailed sequence of causes and biological events leading to disease, whether to address the challenges of delivering personalized medicine, or to identify environmental risks and susceptible subpopulations.

The NexGen framework outlined in Section 2 provides not only categorization of assessments for different applications, but also a process for a controlled and scientifically sound transition from traditional assessment methodology to more advanced technologies as we gain experience and confidence in their use (see Box 11).

Methods illustrated in the Tier 1 and Tier 2 prototypes originally were designed for qualitative evaluation of chemicals. Already, however, some of the approaches are being tested for developing relative potency estimates and quantitative toxicity values (or newer metrics) for use in certain decision contexts. These methods will be used more extensively as they are further developed, and as confidence in the values increases. Importantly, the criteria and scientific process used to evaluate confidence in the new data and application results will guide additional research and further refinement (e.g., focus on hypothesis testing, statistical validity, comparison with real-world values, transparency, peer review, stakeholder communications, and the like).

The Tier 3 data types (1) demonstrate that new methods can provide similar estimates of hazard and risk when compared with results based on traditional data; (2) illustrate the relationships of molecular events to intermediate effects to adverse effects (hazard identification); (3) show how new data can be used to inform exposure-dose-response; and (4) provide a basis for characterizing data-limited chemicals using HT and HC data and adverse outcome pathways (AOPs). Additionally, the prototypes collectively show how to address long-standing risk assessment issues, such as characterizing human variability, assessing cumulative risks, and estimating the quantitative low exposure-response relationships.

Data quality and reporting are significant issues going forward. The searches for data to develop the prototypes resulted in many studies in the literature that could not be used because either the data or the reporting did not meet the criteria for use in health risk assessment. This in part results from the rapid evolution of best practices (i.e., the lag time before being fully implemented in the research community), inconsistent application of criteria for data quality and reporting (see Functional Genomics Data Society 2014 for discussion), and the need for additional guidance and consensus on best practices.

Rhomberg et al. (2013) reviewed 50 existing "weight-of-evidence" frameworks (now termed evidence integration). They identified four phases of analysis consistently used in the 50 frameworks: "(1) defining the causal question and developing criteria for study selection, (2) developing and applying criteria for review of individual studies, (3) evaluating and integrating evidence and (4) drawing conclusions based on inferences" (Rhomberg et al. 2013). Steps 1 and 2, as used at EPA, are discussed in some detail in NRC (2014) and U.S. DHHS (2014). Table 11 focuses on Steps 3 and 4 to evaluate the strength of the causal connections among the exposures, AOP networks, and adverse outcomes discussed in the prototypes and draws on previous authoritative works for the basis of the evidence integration (EPA 2005; Hill 1965; Meek et al. 2014; U.S. DHHS 2014). This table is illustrative; in future practice, evidence integration and conclusions could differ. As presented in Table 11, confidence in causality ranges from suggestive to likely, based on the supporting new data types. "Likely" is generally for cases where the new data types are well anchored to adverse outcomes by a combination of observational and experimental chemical-specific data, similar chemicals data, AOP networks, and robust systems biology understanding. In practice, most new data are anticipated to be suggestive. Of note is that, contrary to traditional approaches, some new approaches can be used to estimate relative potencies or toxicity values in the absence of clearly identified hazards.

The goal of NexGen health assessments, as illustrated by the prototypes, is to improve our understanding of environmental hazards and the environmental concentrations at which those hazards might occur in a population. New types of assessments can be more efficient, and, in some cases, more robust, than those based only on traditional data. Introduction of new assessment types will be iterative and will require input from both the scientific community and the public. Major assessments likely will continue to be driven, for the foreseeable future, by traditional data, however, increasingly augmented with new data types. Concurrently, methods and data for screening and prioritization to support limited scope decision-making will become more prevalent.

Box 11. Applications for New Data Types (adapted from Afshari et al. 2011)

- Elucidation of mechanisms of action
- Classification of compounds by elicited toxicant phenotype
- Generation of hypotheses regarding compound action
- Classification of compounds in similar mechanistic classes
- Ranking and categorization by toxicogenomic signature
- Classification of compounds of unknown toxicity
- Discerning the lowest effect levels for transcripts or BMDs
- Discovery of biomarkers of exposure and toxicity
- Validation/quantification of biomarker signatures
- Discerning dose relationships at environmental exposure levels

Table 11. Illustrative Framework for Causal Determination Focusing on New Data Types

	Prototypes	Evidence for Causality	Evidence Integration
Tier 3	Test the hypothesis that new data types can provide comparable results to traditional data. • Illustrated that new data types (when properly collected, analyzed, and reported) can provide results comparable to those from robust traditional human data. • Indicated that new data types could be used to: (1) evaluate potential hazard of chemicals with no or limited traditional data, (2) augment traditional assessments, or (3) better inform traditional risk assessment issues, such as human variability and susceptibility, cumulative risk, and low exposure-dose-response relationships.	Evidence is consistent, coherent, and biologically plausible that the observed molecular events are causally related to adverse effects. Specifically, in molecular epidemiology and clinical studies: • Specific pattern alterations in molecular events are consistently and strongly associated with known intermediate events and known hazards at environmental exposure levels. • Dose-dependent alterations observed in concomitantly collected molecular events and adverse effects are in the range of environmental exposure of measured exposure-dose relationships (benzene, ozone); PAH exposures were self-reported, and uncertainty in PAH exposures prevented characterization of dose-response. • Adverse outcome pathway (AOP) networks are also disrupted by other chemical and nonchemical stressors known to alter incidence of the specific disease/disorder under consideration (benzene, ozone). • Experimental evidence (pharmacological interventions) has been shown to modify identified AOPs, and have associated an altered incidence of adverse outcomes or severity of disorder (benzene, ozone). • Additional experimental evidence provided by identification of naturally occurring human gene variants in the AOP network that alter susceptibility and risks (benzene, ozone). • Multiple supporting molecular epidemiology and clinical studies; coherent with other systems biology data. See NIH BioSystems (2014a) Acute Myeloid Leukemia or Acute Inflammatory Response. • Data collection, analyses, and reporting met minimum data requirements.	**Implications based on comparisons to robust traditional risk assessments:** For benzene and ozone, identified molecular events are likely causally related to known adverse outcomes in a dose-dependent fashion. The molecular data for PAH are suggestive for a causal association between PAH and lung cancer. Uncertainties in species-to-species extrapolation, and data quality, analysis, and reporting limitations for BaP-associated rodent liver cancer prevented interpretation of BaP molecular data. **Suggestive vs. likely:** More commonly, molecular data are expected to be only suggestive or inadequate for causal determination. To rise to likely, the following are currently necessary: multiple, consistent, high-quality observational studies (across multiple labs/studies); experimental evidence showing that reversal of pathway alterations blocks or ameliorates adverse outcome; or naturally occurring experiments where gene variants alter incidence or characteristic of disease. Important variables such as experimental paradigm (e.g., *in vivo* vs. *in vitro*), cell type, tissue type, and species also require consideration. New data types are likely to be most useful for screening and prioritization, nonregulatory decision-making, and augmenting traditional data, particularly in informing mechanisms of action.

Modification of the Bradford-Hill criteria (consistency, strength, specificity, temporal relationship, and coherence of the data) continued to be useful in the evaluation of data (EPA 2005, 2013c, e; Meek et al. 2014; U.S. DHHS 2014). To simplify the presentation, similar prototypes with shared attributes are aggregated where possible. The left column summarizes the prototype results, the middle column presents evidence for causality exemplified by the prototypes, and the right column illustrates how such prototypic evidence might be integrated and weighed. The first set of prototypes is unique in that the prototypes have known human health effects and well-documented public health risks. For these prototypes, the "Evidence Integration" column evaluates how successful new data types were in predicting known outcomes.

Table 11. Illustrative Framework for Causal Determination Focusing on New Data Types (continued)

	Prototypes	Evidence for Causality	Evidence Integration
Tier 2	Illustrated how large NIH and other large, searchable databases can be knowledge-mined to identify, organize, integrate, and analyze existing data in new ways to discover new insights into public health risks.	• Knowledge mining and meta-analysis discover associations between known exposures to several chemicals (biomonitoring) with prediabetes/diabetes using the Centers for Disease Control and Prevention's National Health Assessment Examination Survey systems • Very limited systems biology context and AOP data.	**Suggestive**: Could rise to likely with the types of supporting data noted above under "Suggestive vs. likely."
	Illustrated how new short-duration *in vivo* exposure bioassays can be used to collect more robust data than *in vitro* exposure bioassays, but in a shorter period than traditional bioassays.	• *In vivo* exposures of intact organisms with intact metabolism associate molecular events with adverse outcomes or measure adverse outcomes directly. • High-content assays with measurable adverse outcomes (e.g., zebrafish developmental assay) have greater evidentiary weight than initiating event assays. • Cross-species extrapolation introduces additional uncertainties. • Relatively well-understood systems biology context and AOP are necessary for interpretation of the data.	• **Suggestive:** For transcriptomic studies with AOP descriptions. This could rise to likely with the types of supporting data noted above under "Suggestive vs. likely." • **Suggestive:** Could rise to likely for human health hazard using zebrafish developmental outcomes and other models with phenotypic outcomes; could rise to likely with the types of supporting data noted above under in "Suggestive vs. likely."
Tier 1	Illustrated recent advances and use of QSAR models to estimate values similar to those of traditionally based assessments.	• QSAR models can predict chemical-specific toxicity values based on chemical inherent properties for a number of data poor chemicals. • Models are developed based on chemical structures and known outcomes for data rich chemicals. • OECD is harmonizing international use of QSAR hazard models and read-across in the OECD QSAR toolbox	• **Suggestive:** TopKat Model predictions of potency when model is appropriate for chemicals evaluated; not generally predictive of dose-response for specific hazards; does generate a LOAEL for a subset of the data poor chemicals that meet confidence criteria. Additional OECD models and read-across can improve confidence in hazard characterization. • Could rise to likely with the types of supporting data noted above under "Suggestive vs. likely" data to adverse outcomes.
	Illustrated how new, generally robotically conducted, *in vitro* bioassays can evaluate (with unprecedented speed) the potential of chemicals to disrupt biological processes.	• High-throughput *in vitro* assays based on biological process disruptions are interpreted in a systems biology and AOP context, and associated with adverse outcomes. • Thyroid hormone disruptor assay results are supported by considerable systems biology and cross-species understanding. See NIH BioSystems (2014a) for additional review of thyroid hormone-mediated signaling pathways.	• **Suggestive:** When coupled with understanding of the AOP(s); could rise to likely with the types of supporting data noted above under "Suggestive vs. likely" data to adverse outcomes

Modification of the Bradford-Hill criteria (consistency, strength, specificity, temporal relationship, and coherence of the data) continued to be useful in the evaluation of data (EPA 2005, 2013c, e; Meek et al. 2014; U.S. DHHS 2014). To simplify the presentation, similar prototypes with shared attributes are aggregated where possible. The left column summarizes the prototype results, the middle column presents evidence for causality exemplified by the prototypes, and the right column illustrates how such prototypic evidence might be integrated and weighed. The first set of prototypes is unique in that the prototypes have known human health effects and well-documented public health risks. For these prototypes, the "Evidence Integration" column evaluates how successful new data types were in predicting known outcomes.

With new approaches we can (1) gather new data on biological alterations caused by chemical exposures; (2) begin to understand AOPs and AOP networks, and improve our interpretation of new data in a biological context; (3) start understanding the effects of other environmental risk factors or modifying factors, such as mixtures exposure, other environmental stressors, and susceptibility factors like genetic makeup and preexisting health status; (4) better characterize exposure-response; and (5) better characterize variability and uncertainties. New data types will support assessments based on an understanding of adverse outcome and the underlying mechanisms needed to identify causal links between exposures and effects. Conversely, the new data can be used to identify network interactions[49] that represent "normal" biology and the chemical perturbations that lead to adverse outcomes (Andersen, M. E. and Krewski 2009; Chiu et al. 2013; Goodman et al. 2014).

5.1 Looking Across the Major-scope Assessment Prototypes (Tier 3)

The Tier 3 prototypes were designed to test the hypothesis that new data types could provide results comparable to those that robust traditional data provide (see Section 3 and EPA 2013c, d; Hatch et al. in press; McCullough et al. in press; McHale et al. 2010; McHale et al. 2012; Smith, M. T. et al. 2011; Thomas, R. et al. 2014). Support for this hypothesis follows.

AOP networks appeared useful in predicting specific hazards and could successfully do so for benzene and other known leukemogens (hematotoxicity); ozone (lung inflammation and injury); and polycyclic aromatic hydrocarbons (PAHs; lung cancer). Nonchemical stressors that alter risks also appear to affect the same AOP networks as chemical risk factors. These exposure-dependent network modifications appear causally related to specific gene changes, pathway perturbations, intermediate events, and adverse effects. We inferred from these data that less well-studied chemicals that induce the same AOP or AOP network would be of concern for the same health outcomes. Thus, AOP networks such as those developed by EPA, the Organization for Economic Cooperation and Development (OECD), or the National Institutes of Health (NIH) BioSystems are anticipated to be essential in the future to help elaborate mechanisms of action and potentially increase confidence in the overall evidence; assess hazards posed by less well-studied chemicals; and provide a construct for grouping chemical and nonchemical stressors by common mechanisms for cumulative assessment. As illustrated by the prototypes, AOP networks also can help evaluate the role of human gene variants in subpopulation susceptibility (or resistance).

An AOP network, or component biomarkers, can help characterize exposure-dose-response relationships, as illustrated by benzene and ozone (and the Tier 2 thyroid hormone disruption prototype discussed below).[50] Important to note is that AOPs appear to evolve with increasing exposures. For example, with benzene, gene and pathway alterations indicative of impaired immune function are present at all exposure levels evaluated (from <0.1 ppm to ≤10 ppm), but at

[49]As noted in earlier in the report, AOPs and AOP networks do not imply creation of new biological processes that are specifically adverse, rather they address perturbations of normal biological processes.
[50]Uncertainty around self-reported PAHs exposures (in the available data sets used here) prevented characterization of exposure-dose-response for PAHs.

higher concentrations AOPs characteristic of more frank toxicity (apoptosis and cell death) begin to emerge. Thus, data collection over a wide range of environmental concentrations remains important for interpreting new data types. One of the most promising applications of exposure/effect biomarkers is the ability to measure directly events of interest in environmentally exposed humans; such applications are revolutionizing epidemiology.

Chemical exposures known to result in specific diseases share AOP networks with disease of unknown origins (idiopathic or potentially naturally occurring disease). Chemically induced adverse effects appear to add to naturally occurring backgrounds of disease, via shared mechanisms. As discussed by NRC (2009) and Crump et al. (1976), this finding has implications for an assumption of low-dose linearity for cancer and noncancer outcomes at the population level.

Uncertainties evaluated, where possible, and that deserve consideration in risk assessment as feasible, arise from the following factors: interindividual and subpopulation variability from genetic makeup and coexposures; species, target versus nontarget cell and tissue types; and *in vivo* versus *in vitro* primary cell culture and cell line protocols.

The evidence for these conclusions is additionally summarized in Table 11, and discussed in more detail in Section 3 and the NexGen background papers.

5.2 Looking Across the Limited-scope Assessment Prototypes (Tier 2)

The Tier 2 approaches appear useful in identifying potential hazards, characterizing relative potency of hundreds of chemicals, and using AOP networks to refine both hazard identification and exposure-response assessment. Two very different approaches were considered in the limited-scope prototypes: (1) computer-assisted knowledge-mining techniques used to scan huge existing databases to identify associations among various factors of interest such as exposure, health status, coexposures, and genetic and lifestyle susceptibility traits (Burgoon 2011; Patel et al. 2012a; Patel et al. 2013); and (2) relatively new experimental paradigms involving short-duration *in vivo* exposure of both alternative (nonmammalian) and mammalian species to predict health outcomes, to explore interactions of AOPs and apparent exposure-response anomalies, and to consider species-to-species similarities and differences (Padilla et al. 2012; Perkins et al. 2013; Skolness et al. 2013; Thomas, R. S. et al. 2012c; Thomas, R. S. et al. 2013b; 2013c; Warner et al. 2012). These new approaches are faster and less expensive than the molecular epidemiology and molecular clinical studies noted above. Furthermore, unlike the quantitative structure activity relationship (QSAR) models and HTS data (discussed below), the data from *in vivo* studies are from intact systems for metabolism, normal architecture (for various cell types), and normal tissue interactions; and can be used to study more complex system-level outcomes, such as developmental and neurobehavioral outcomes. Confidence in these data generally ranks between Tier 3 and Tier 1 approaches. Highlights from the prototypes are briefly discussed below.

Computer algorithms were developed to search the NHANEs database and identify associations between chemical exposures and incidence of prediabetes or diabetes. Exposures were determined via the National Health and Nutrition Examination Survey (NHANES) human tissue biomonitoring; incidence was clinically defined within NHANES. In all four data-mining exercises, specific chemical exposures were associated with altered diabetes or prediabetes risks (e.g., chlorinated organics, heavy metals, selected nutrients). Because data mining identifies associations among events in very

large data sets, results are most suitable for hypothesis generation. The addition of other data types, such as AOP network data, read-across, or traditional data, augment confidence in the observed associations. Thayer et al. (2012) reported on a workshop that reviewed traditional data on chemically related diabetes and obesity, and independently identified a similar set of chemicals to those identified in the above data mining exercises.

Two Tier 2 prototypes demonstrated use of short-duration exposures in alternative species and mammalian species, respectively, coupled with new molecular and computational approaches to provide insights into potential environmental risks. The alternative species assays were used to detect effects over the entire lifespan of the organism, and to evaluate population dynamics. The mammalian assay assumed that molecular events identified in short-duration experiments would reflect chronic outcomes and thus be useful in more rapid assessment of chemicals. These short-duration exposure studies using different animal models successfully identified exposures associated with molecular events, AOPs, and AOP networks; explored complex mechanistic behaviors; screened for potential hazards; and evaluated chemical potencies.

Although only one prototype illustrated data-mining approaches, data-mining is becoming an essential tool in many areas of modern science and in the development of assessments in all tiers. With the explosive growth of new data, evaluation of the available literature rarely can be accomplished without using computer algorithms to search for, identify, organize, prioritize, and integrate key data.

5.3 Looking Across the Prioritization and Screening Prototypes (Tier 1)

For the first time in the history of risk assessment, robotically conducted, *in vitro* experiments are allowing the evaluation of chemicals (e.g., on the order of 10,000). Results from QSAR models (Goldsmith et al. 2012; Venkatapathy and Wang 2013; 2012b; Wang, N. et al. 2012c) and HT *in vitro* bioassays were used to illustrate a set of methods to evaluate chemicals rapidly (Judson et al. 2013; Kavlock et al. 2012; Rusyn et al. 2012; Sipes et al. 2013; Tice et al. 2013). Kavlock et al. (2012) note that "These tools can probe chemical-biological interactions at fundamental levels, focusing on the molecular and cellular pathways that are targets of chemical disruption."

Thousands of chemicals are currently being evaluated in the ToxCast and Tox21 programs using these methods. Estimates of relative potency and insights on potential hazards are being generated.

Methods are being developed using reverse dosimetry to extrapolate *in vitro* concentration to test species (e.g., rodent) and human *in vivo* concentrations (*in vitro*-to-*in vivo* extrapolation [IVIVE]; see Section 3.3.2.3) (Hubal 2009; Rotroff et al. 2010; Wetmore et al. 2013; Wetmore et al. 2012). This extrapolation supports quantitative comparisons of *in vitro* toxicity results with *in vivo* results and estimates of dose-response for human exposures.

With the current state of the science, estimates of risks of disease in humans based exclusively on *in vitro* findings are too uncertain, and are primarily useful for screening and ranking large numbers of chemicals for further evaluation and assessment. Insights on

underlying mechanisms of toxicity, and the factors that might contribute to the variability in response to chemical exposure, however, are progressing from these data streams and increasing their utility in understanding risks (Lock et al. 2012).

5.4 Certain Caveats Pertaining to New Data Types in Risk Assessment

In general, much of the new toxicogenomic data currently being generated is associative in nature, that is, exposure and adverse outcomes can be associated with hundreds to thousands of gene changes, not all of which are likely to be causal in nature (Mendrick 2011). Associative data are only "suggestive" of a causal relationship between exposure and adverse health outcomes. Criteria to move from "suggestive" to "likely" causal include meta-analyses of multiple, independent studies yielding similar results, experimental evidence of alterations in putative AOP networks with consequent health outcomes (such as pharmacological interventions, gene knock-in/-out studies, or alterations in risks due to human gene variants in key pathways), or combinations of traditional and NexGen data. The prototypes demonstrated how different types of evidence in each decision support category might be characterized with respect to causality and evidence integration. This is shown in Table 11. Additionally,

- Cell type, tissue, individual, subpopulation, species, and test system can affect how specific alterations in molecular events manifest as adverse outcomes or disease, even when the molecular signature is the same. This phenomenon is likely due, at least in part, to epigenomic differences and genomic plasticity. This issue should be considered within an assessment, as is feasible.
- The metabolism of many chemicals often plays an important role in toxicity. That most HT *in vitro* test systems are not metabolically competent should be taken into account. Although various approaches to add metabolic capability are being evaluated, satisfactory solutions are not yet available. Consequently, positive results can be informative, but negative results should not be interpreted as lack of toxicity.
- Molecular profiles appear time-dependent, that is, they evolve over time with continued exposu*re and p*ost-exposure. Predicting adverse outcomes therefore can be challenging based only on "snapshots" of biological events. Some signatures do appear to be stable over time, and might serve as reliable indicators of chronic outcomes.
- Adverse outcome arguments in support of a regulatory assessment cannot be made solely with gene expression data, as messenger ribonucleic acid (mRNA) expression levels cannot be used to infer protein activity directly. These data could, however, be suitable for ranking and screening. Gene expression data can also be used in a regulatory assessment to complement other mechanistic data.
- Data reproducibility and false negative rates remain potential limitations of HTS/HCS assays (e.g., toxicogenomics). The false negative rate (i.e., deeming a chemical nontoxic when it is toxic) tends to decrease as the number of independent replicates used increases. Successful screening programs require low false negative rates, while balancing their efficiencies (i.e., cost and throughput).

- Our current ability to monitor multiple molecular processes (i.e., genomics, transcriptomics, proteomics, and epigenomics) in a single study is very limited, primarily due to expense. This lack of biological integration limits our understanding.
- Sufficient good-quality data from the open literature adequate to support risk assessment are available for a limited number of chemicals, at this time, due primarily to experimental design and reporting issues. This lack of data underscores the critical importance of high-quality research and testing programs like ToxCast and Tox21 to advance the methods development; it also emphasizes the need for systematic review of the data.

5.5 Fit-for-purpose Assessment

Table 12 integrates many of the lessons learned from the NexGen effort and illustrates components of "fit-for-purpose" assessments matched to the decision-context categories. Listed in the table are potential uses for NexGen assessments, data sources and types in different assessment categories, exposure paradigms used, incorporation of metabolism and toxicokinetics, use of traditional data, hazard characterization, potency metrics, inferences drawn about the causal associations among exposure, AOPs and adverse outcomes, and the numbers of chemicals that can be assessed over a given time period.

5.6 Conclusions

Based on the lessons learned in the NexGen program, several new types of high- and medium-throughput assessments are being advanced. In the foreseeable future:

- Tens of thousands of chemicals with no or very limited traditional data will be analyzed using similarities in physical-chemical structure of known toxicants to estimate the toxicity of unstudied chemicals (often called quantitative structure-activity modeling); and using rapid, robotically conducted *in vitro* bioassay data to identify a chemical's potency to alter important biological processes as indicators of toxicity (e.g., ToxCast and Tox21 programs).
- Thousands of chemicals will be evaluated using computer-driven analyses of the world's new and existing data, extracted from the published literature and stored in massive databases, to develop new knowledge about the potential toxicity of chemicals, and the causes of disease. Examples of such databases include the National Library of Medicine's National Center for Computational Biology databases and the Comparative Toxicogenomic Database (CTD). Previously, analyzing so much data from so many sources in an integrated fashion was not possible.
- Hundreds of chemicals will be evaluated using a variety of new methods, including a concerted, mechanistic approaches to understanding the cumulative effects posed by multiple chemical and nonchemical stressors.

Issues of particular interest, likely to be informed by new and emerging knowledge, are historically difficult risk assessment questions such as: Why do individuals and specific populations respond differently to environmental exposures? Are children at particular risk for certain exposures and effects? What happens when people are exposed to low levels of many chemicals? How might other environmental factors like poverty and preexisting ill health make chemical exposures riskier?

Table 12. Illustrative Fit-for-Purpose Assessments Matched to the Decision-context Categories

Description	Tier 1 Prioritization and Screening	Tier 2 Limited scope Assessments	Tier 3 Major scope Assessments
Potential Uses of NexGen Assessments	• Screening chemicals with no-data-other-than-QSAR-or-HT-data • Queuing for research, testing, or assessment • Urgent or emergency response	Generally nonregulatory decision-making • Urban air toxics • Potential water contaminants • Hazardous waste and superfund chemicals • Urgent or emergency response	Often regulatory decision-making • National risk assessments • Community risk assessment • Special problems
Data Sources	EPA databases such as ACToR and ToxCast	NIH databases, Array express, NHANES	All policy-relevant data
New Data Types[a]	QSAR, high-throughput screening, read across	High-content assays, medium throughput assays, knowledge mined large data sets, AOPs	Molecular epidemiology, clinical and animal studies
Exposure Paradigms of Studies Used in Assessments	*In vitro,* in *silico*	*In vitro, in situ,* and *in vivo, in silico*	*In vivo*
Metabolism in Test Systems	Little to none	Partial to intact	Intact
Incorporation of Toxicokinetics	Reverse toxicokinetic models	Reverse toxicokinetics models, biomonitoring	Dosimetry and PK modeling
Traditional *In Vivo* Data	Anchors *in vitro* assays using pesticide registration data	No to very limited	New data types augment traditional data that remain basis for assessment
Hazards	Nonspecific	Nonspecific to Identified	Identified
Potency Metrics	Relative rankings and toxicity values	Relative rankings and toxicity values	Risk distributions, cumulative risks, community risks
Strength of Evidence Linking Exposure to Adverse Effects	Suggestive	Suggestive to likely	Suggestive to known
Numbers of Chemicals that Can Be Assessed	1000s–10,000s	100s–1000s	100s
Time to Conduct Assessment	Hours–Days	Hours–Weeks	Days–Years

[a]Each assessment type also uses the data types from the column to the left.

Such large-scale knowledge creation was unimaginable 15 years ago. This new knowledge holds great promise for improving our ability to conduct risk assessments, and to protect human health and the environment.

Logistical and methodological challenges in interpreting and using newer data and methods in risk assessment remain significant. Despite these challenges, we anticipate that these new approaches will have a variety of applications for risk managers within EPA and the risk assessment community at large in the near future. Such applications include identifying safer chemicals and processes, and reducing hazardous chemicals in the environment. Near-term progress will include case-by-case

development of additional examples made available for public input and peer review. The research implications generated from this report are captured in EPA's Chemical Safety for Sustainability (CSS) and Human Health Risk Assessment (HHRA) research program plan, and the National Institute of Environmental Health Sciences' (NIEHS) Strategic Plan. EPA's research plans are discussed in more detail in Section 6.

6 Challenges and Research Directions

6.1 Challenges

More than 80,000 chemicals are currently listed or registered for use in the United States under EPA authorities, and at least a thousand more are introduced every year (EPA 2014f). The overarching challenge is to obtain and interpret data that provide the information risk assessors need to assess these chemicals quickly and efficiently for safety and sustainability. The information needed includes the following: (1) how best to design and produce safer chemicals, (2) how chemicals and their byproducts move through the environment, (3) what the sources of chemical exposure are, (4) what are the critical biological processes and toxicity pathways that chemicals might interact with to cause disease, and (5) what is the contribution of exposure to chemicals in the environment to the overall disease burden for susceptible populations (EPA 2012b).

The prototypes presented in this report demonstrate how new data types (molecular, cellular, tissue, whole body) can be used to address (1), (4), and (5) above. Arguably, the greatest challenge to the use of molecular data in risk assessment is interpreting those data to predict observable adverse effects in humans. In other words, how do changes in molecular events affect cells, changes in cells affect tissues and organs, and changes in organs affect the whole body? Large amounts of HTS/HCS data are being collected on effects at the molecular level, and the body of information on diseases and disease outcomes is substantial, yet only very sparse data are available on intermediate levels of organization and on the sequence of events from disruption of normal biology at the cell level to effects at higher levels of organization.

To fill these gaps in our understanding of the complex chemical and biological interactions at different levels of biological organization, advanced research programs and models are needed. Specific areas of interest include the following:

- reliable, predictive molecular indicators for a wide variety of chemicals and diseases to assess hazard and characterize exposure-dose-response;
- identification of the networked interactions among genes, proteins, cells, tissues, organs, individuals and populations; and the sequence of events at different levels that can lead to disease (i.e., adverse outcome pathway (AOP) networks; Hartung and McBride 2011);
- an integrated understanding of how genes are expressed, and how the resulting proteins interact to maintain the body;
- methods to group chemical and nonchemical stressors based on common AOPs to enable cumulative risk assessment;
- methods to measure and account for individual human variability due to genetic differences, preexisting backgrounds of disease and exposure, or adaptive and compensatory capabilities; and how to incorporate this information to assess risk at the population level;
- data and methods to adjust for interspecies differences when assessing potential toxicity in humans based on nonhuman toxicity data; and

- data and methods to characterize the dose-response curve quantitatively for responses at low levels of exposure.

Verifying high throughput/high content toxicity testing schemes and computational models is essential for these new data and approaches to be used for risk-based decisions or in risk assessments. Central to this effort is a framework and criteria for determining the adequacy of the new data types for different types of decisions. The level of certainty needed in the data varies with its use because inaccuracies in results have increasing consequence and costs as one progresses from decisions about screening, to further testing, to what are safe levels, to what regulatory actions need to be taken (Crawford-Brown 2013). Traditional "validation" schemes designed to evaluate conventional assay and testing structures do not adequately address the potential uses of these new data and methods, and would require an impractical number of years to implement. Thus, as the technology for rapid, efficient, robust hazard testing advances, the verification process for these new methods must also advance to provide confidence in their use. Clear and transparent articulation of these decision considerations will be important to the acceptance of, and support for, assessment results.

6.2 Research Directions

EPA's Office of Research and Development (ORD; EPA 2014c) has the lead on identifying and conducting EPA research to address the above challenges. ORD has six national research programs, two of which are discussed here that directly address innovation and development of NexGen risk assessments: (1) the Chemical Safety for Sustainability (CSS; EPA 2014a) research program; and (2) the Human Health Risk Assessment (HHRA; EPA 2014b) research program. The discrepancy in available data across levels of biological organization and over time is a major focus of ongoing research in both programs. CSS develops new tools and innovative technologies to evaluate chemical toxicity, to optimize confidence in risk management decisions, and to prioritize time-critical research. HHRA incorporates and integrates the available tools and scientific information into state-of-the-science risk assessments that support regulatory actions to protect human health and the environment.

Insights gained during the development of the prototypes presented in this report (see Section 5) are guiding further research. Specific areas of focus are reflected in the top level CSS and HHRA research themes and areas of interest bulleted below. EPA freely provides the details of the strategic research action plans in the CSS and HHRA programs (2012b, d). EPA also collaborates with numerous other research centers. Appendix A briefly summarizes relevant research activities with EPA's collaborators in the United States and in Europe (where complementary, equally compelling research is underway) to advance the next generation of toxicity testing and risk assessment. Highlights of ongoing research sponsored by NIEHS (2014c) is also listed below.

Top Themes in EPA's Ongoing Chemical Safety and Sustainability Research Program (EPA 2012b) include the following:

- Sustainable Chemistry;
- High-Throughput Toxicity Assay Development, Predictive Models, Integrated Testing Strategies;
- Rapid Exposure and Dosimetry Tools and Data;
- Evaluation of Alternative Assays and Applications in Hazard Assessment;
- Chemical Evaluation for Emerging Materials;
- Life Cycle and Human Exposure Modeling;
- Integrated Modeling for Ecological Risk Assessment;
- AOP Discovery and Development;
- Systems Biology Computational Models – Virtual Tissue (VT) Models based on Advanced *In Vitro* (e.g., organotypic systems), Alternative Species *In Vivo* Data, and Knowledge Mining; and
- Integrated Applications and User Interfaces to Support Decision-making.

Top areas in EPA's Human Health Risk Assessment Research Program (EPA 2012d) include:

- Identify, evaluate, integrate, and apply relevant data from a variety of scientific disciplines to characterize the risk from exposures of individual chemicals, mixtures and nonchemical stressors.
- Develop a suite of state-of-the-science assessment products that inform a variety of risk-based decisions by the EPA, State/local/tribal agencies and the public to protect public health and the environment (e.g., ISAs, IRIS, MSDs, PPRTVs[51]).
- Broaden exposure assessment technology and assessment guidance to translate exposure and dose estimates across various experimental designs to address different exposure scenarios flexibly.
- Update dosimetry modeling and biomarker approaches to predict a profile of internal dose metrics across all routes to support mode of action (MOA)/AOP, and aggregate or cumulative risk descriptions.
- Expand cumulative risk assessment methods to incorporate ecological impacts and indices of resilience and wellness to support sustainability and community risk characterizations.
- Improve prioritization and emergency response by evaluating and incorporating new data streams, and developing rapid assessment approaches.

[51]ISA = Integrated Science Assessments for six principal pollutants - ozone, particulate matter, carbon monoxide, sulfur dioxides, nitrogen oxides, and lead; IRIS = Integrated Risk Information System human health assessments on more than 550 chemical substances; MSD = Multipollutant Science Documents; PPRTV = Provisional Peer-Reviewed Toxicity Value.

- Advance decision analytic and probabilistic approaches to characterize response functions more fully and better inform cost-benefit analyses.
- Enhance data access and management systems to support transparency and efficiency.
- Develop and apply effective methods for stakeholder engagement and risk assessment training to varied audiences through the Risk Assessment Training and Experience (RATE) program.

Highlights of NIEHS-sponsored research – Mapping the Human Toxome by Systems Toxicology (NIEHS 2014c) include:

- Comprehensively map pathways of endocrine disruption as a first step toward mapping the human toxome (the entirety of pathways of toxicity in humans).
- Leverage rapidly evolving scientific understanding of how genes, proteins, and small molecules interact to form molecular pathways that maintain cell function, applying orthogonal omics approaches (transcriptomics, metabolomics) to map and annotate toxicity pathways for a defined set of endocrine disruptors.
- Conduct a series of stakeholder workshops to enable development of a consensus-driven process for pathway annotation, validation, sharing and risk assessment.
- Develop a public database on toxicity pathways, providing a common, community-accessible framework that will enable the toxicology community at large to map the human toxome comprehensively and cooperatively using integrated testing strategies.
- Verify the identified pathways of toxicity, and extend the concepts to additional toxicants, cell systems, and endocrine disruptor hazards to additional omics platforms and to dose response modeling.

ORD will continue to elaborate the NexGen framework, identify hazards posed by environmental factors, estimate potencies of toxic chemicals to cause harm, and characterize risk to the general population and sensitive subpopulation. These efforts will incorporate the information from new biology targeted to specific risk assessment purposes. ORD also will work with EPA's Program Offices using Tier 1 screening and prioritization approaches to queue up new assessments. Results from this work will be used to refine the testing paradigm and inform research.

Toxicity values informed by new types of knowledge will be developed in each tier and decision context, from needs to screen chemicals for future testing to the development of reference values for a larger number of chemicals. Levels of confidence in those values will be characterized depending on the types and quality of the supporting data. Examples will be identified where molecular (and higher level) biology data might be considered for Tier 3 assessments to augment traditional assessment methodologies. These examples will provide more opportunities to solicit public comment and peer review. A verification process will be developed for new methods and data types with a focus on clear articulation of the considerations for incorporating results into different decision contexts and into the overall integration of evidence for a risk assessment. The goal will be to increase confidence in assessments that include these new approaches. Significant

scientific gaps will continue to be identified from ongoing prototype development, and addressed in future research planning.

Logistical and methodological challenges in interpreting and using newer data and methods in risk assessment remain significant. Despite these challenges, we anticipate that the new approaches demonstrated in the prototypes will have a variety of applications for risk managers within EPA and the risk assessment community at large in the near future, including identifying safer chemicals and processes and reducing risk from exposures to hazardous chemicals in the environment. Near-term progress will include case-by-case development of additional examples that are made available for public input and peer review. The research gaps identified in this report will continue to guide research at EPA and throughout the world. The reader is encouraged to frequent the internet sites of EPA and other research programs to learn about the latest developments and progress toward planned objectives in this rapidly evolving science.

7 References

Abdo N, Xia M, Kosyk O, Huang R, Sakamuru S, Austin CP, et al. (2012). The 1000 genomes toxicity screening project: Utilizing the power of human genome variation for population-scale *in vitro* testing. *51st Annual Meeting of the Society of Toxicology* Poster Abstract #765. http://www.toxicology.org/AI/Pub/Tox/2012Tox.pdf.

Abecasis GR, Auton A, Brooks LD, DePristo MA, Durbin RM, Handsaker RE, et al. (2012). An integrated map of genetic variation from 1,092 human genomes. *Nature* 491: 56-65. http://www.ncbi.nlm.nih.gov/pubmed/23128226.

Afridi HI, Kazi TG, Kazi N, Jamali MK, Arain MB, Jalbani N, et al. (2008). Evaluation of status of toxic metals in biological samples of diabetes mellitus patients. *Diabetes Research & Clinical Practice* 80: 280-288. http://www.ncbi.nlm.nih.gov/pubmed/18276029.

Afshari CA, Hamadeh HK, Bushel PR. (2011). The evolution of bioinformatics in toxicology: Advancing toxicogenomics. *Toxicological Sciences* 120: S225-S237. http://toxsci.oxfordjournals.org/content/120/suppl_1/S225.full.pdf+html.

Agrawal R, Imielinski T, Swami A. (1993). Mining association rules between sets of items in large databases. *Proceedings of the 1993 ACM SIGMOD International Conference on Management of Data* 22: 207-216. http://www.almaden.ibm.com/cs/quest/papers/sigmod93.pdf.

Aldenberg T, Rorije E. (2013). Species Sensitivity Distribution estimation from uncertain (QSAR-based) effects data. *Altern Lab Anim* 41: 19-31. http://www.ncbi.nlm.nih.gov/pubmed/23614542.

Alexanian A, Sorokin A. (2013). Targeting 20-HETE producing enzymes in cancer - rationale, pharmacology, and clinical potential. *Onco Targets Ther* 6: 243-255. http://www.ncbi.nlm.nih.gov/pubmed/23569388.

Andersen ME, Krewski D. (2009). Toxicity testing in the 21st century: Bringing the vision to life. *Toxicol Sci* 107: 324-330. http://www.ncbi.nlm.nih.gov/pubmed/19074763.

Andersen S, Pedersen KM, Bruun NH, Laurberg P. (2002). Narrow individual variations in serum T(4) and T(3) in normal subjects: A clue to the understanding of subclinical thyroid disease. *Journal of Clinical Endocrinology & Metabolism* 87: 1068-1072. http://www.ncbi.nlm.nih.gov/pubmed/11889165.

Angerer J, Bird MG, Burke TA, Doerrer NG, Needham L, Robison SH, et al. (2006). Strategic biomonitoring initiatives: Moving the science forward. *Toxicological Sciences* 93: 3-10. http://www.ncbi.nlm.nih.gov/pubmed/16785253.

Ankley GT, Bennett RS, Erickson RJ, Hoff DJ, Hornung MW, Johnson RD, et al. (2010). Adverse outcome pathways: A conceptual framework to support ecotoxicology research and risk assessment. *Environmental Toxicology and Chemistry* 29: 730-741. http://service004.hpc.ncsu.edu/toxicology/websites/journalclub/linked_files/Fall10/Environ%20Toxicol%20Chem%202010%20Ankley.pdf.

Ankley GT, Gray LE. (2013). Cross-species conservation of endocrine pathways: A critical analysis of tier 1 fish and rat screening assays with 12 model chemicals. *Environmental Toxicology and Chemistry* http://www.ncbi.nlm.nih.gov/pubmed/23401061.

Arnot JA, Mackay D TU Canadian Environmental Modelling Centre. (2007). *Risk Prioritization for a Subset of Domestic Substance List Chemicals Using the RAIDAR Model*. Retrieved from http://www.trentu.ca/academic/aminss/envmodel/CEMC200703.pdf.

Arnot JA, Mackay D, Parkerton TF, Zaleski RT, Warren CS. (2010a). Multimedia modeling of human exposure to chemical substances: The roles of food web biomagnification and biotransformation. *Environmental Toxicology and Chemistry* 29: 45-55. http://www.ncbi.nlm.nih.gov/pubmed/20821418.

Arnot JA, Mackay D, Sutcliffe R, Lo B. (2010b). Estimating farfield organic chemical exposures, intake rates and intake fractions to human age classes. *Environmental Modelling & Software* 25: 1166-1175. http://www.sciencedirect.com/science/article/pii/S136481521000085X.

ATSDR (Agency for Toxic Substances and Disease Registry). (1995). *Toxicological Profile for Polycyclic Aromatic Hydrocarbons (PAHs)*. Retrieved from http://www.atsdr.cdc.gov/toxprofiles/TP.asp?id=122&tid=25#bookmark06 (accessed August 29, 2014).

ATSDR (Agency for Toxic Substances and Disease Registry). (2007). *Toxicological profile for benzene*. Retrieved from http://www.atsdr.cdc.gov/toxprofiles/tp3.pdf (accessed August 29, 2014).

Attene-Ramos MS, Miller N, Huang R, Michael S, Itkin M, Kavlock RJ, et al. (2013). The Tox21 robotic platform for the assessment of environmental chemicals - from vision to reality. *Drug Discov Today* 18: 716-723. http://www.ncbi.nlm.nih.gov/pubmed/23732176.

Auerbach SS, Shah RR, Mav D, Smith CS, Walker NJ, Vallant MK, et al. (2010). Predicting the hepatocarcinogenic potential of alkenylbenzene flavoring agents using toxicogenomics and machine learning. *Toxicology and Applied Pharmacology* 243: 300-314. http://dx.doi.org/10.1016/j.taap.2009.11.021.

Baker M. (2010). Epigenome: Mapping in motion. *Nature Methods* 7: 181-186. http://www.nature.com/nmeth/journal/v7/n3/abs/nmeth0310-181.html.

Ball N, Bartels M, Budinsky R, Klapacz J, Hays S, Kirman C, et al. (2014). The challenge of using read-across within the EU REACH regulatory framework; how much uncertainty is too much? Dipropylene glycol methyl ether acetate, an exemplary case study. *Regulatory Toxicology and Pharmacology* 68: 212-221. http://www.sciencedirect.com/science/article/pii/S0273230013002225.

Barton HA, Chiu WA, Woodrow Setzer R, Andersen ME, Bailer AJ, Bois FY, et al. (2007). Characterizing uncertainty and variability in physiologically based pharmacokinetic models: State of the science and needs for research and implementation. *Toxicological Sciences* 99: 395-402. http://www.ncbi.nlm.nih.gov/pubmed/17483121.

Basketter DA, Clewell H, Kimber I, Rossi A, Blaauboer B, Burrier R, et al. (2012). A roadmap for the development of alternative (non-animal) methods for systemic toxicity testing - t4 report*. *Alternatives to Animal Experiments* 29: 3-91. http://www.ncbi.nlm.nih.gov/pubmed/22307314.

Bauer AK, Travis EL, Malhotra SS, Rondini EA, Walker C, Cho HY, et al. (2010). Identification of novel susceptibility genes in ozone-induced inflammation in mice. *Eur Respir J* 36: 428-437. http://www.ncbi.nlm.nih.gov/pubmed/20032013.

Behl M, Rao D, Aagaard K, Davidson TL, Levin ED, Slotkin TA, et al. (2013). Evaluation of the association between maternal smoking, childhood obesity, and metabolic disorders: A national toxicology program workshop review. *Environmental Health Perspectives* 121: 170-180. http://www.ncbi.nlm.nih.gov/pubmed/23232494.

Béliveau M, Tardif R, Krishnan K. (2003). Quantitative structure-property relationships for physiologically based pharmacokinetic modeling of volatile organic chemicals in rats. *Toxicology and Applied Pharmacology* 189: 221-232. http://www.ncbi.nlm.nih.gov/pubmed/12791307.

Bell R, Early JL, Nonavinakere VK, Mallory Z. (1990). Effect of cadmium on blood glucose level in the rat. *Toxicology Letters* 54: 199-205. http://www.ncbi.nlm.nih.gov/pubmed/2260118.

Bell S, Edwards S. (2014). Data mining NHANES to identify environmental chemical and disease associations.Proceedings of the 2014 SIAM International Conferecnce on Data Mining.

Belson M, Kingsley B, Holmes A. (2007). Risk factors for acute leukemia in children: A review. *Environmental Health Perspectives* 115: 138-145. http://www.ncbi.nlm.nih.gov/pubmed/17366834.

Benigni R. (2013). Evaluation of the toxicity forecasting capability of EPA's ToxCast Phase I data: Can ToxCast *in vitro* assays predict carcinogenicity? *Journal of Environmental Science and Health, Part C* 31: 201-212. http://dx.doi.org/10.1080/10590501.2013.824188.

Berg EL, Yang J, Melrose J, Nguyen D, Privat S, Rosler E, et al. (2010). Chemical target and pathway toxicity mechanisms defined in primary human cell systems. *Journal of Pharmacological and Toxicological Methods* 61: 3-15. http://www.ncbi.nlm.nih.gov/pubmed/19879948.

Bhattacharya S, Zhang Q, Carmichael PL, Boekelheide K, Andersen ME. (2011). Toxicity testing in the 21 century: Defining new risk assessment approaches based on perturbation of intracellular toxicity pathways. *Public Library of Science One* 6: e20887. http://www.ncbi.nlm.nih.gov/pubmed/21701582.

Birnbaum LS. (2012). NIEHS and the future of toxicology. Presentation. *FutureTox Meeting of the Society of Toxicology* October 18-19. Arlington, VA.

Blackburn K, Stuard S. (2014). A framework to facilitate consistent characterization of read across uncertainty. *Regulatory Toxicology and Pharmacology* 68: 353-362. http://www.ncbi.nlm.nih.gov/pubmed/24457134.

Blakesley V, Awni W, Locke C, Ludden T, Granneman GR, Braverman LE. (2004). Are bioequivalence studies of levothyroxine sodium formulations in euthyroid volunteers reliable? *Thyroid* 14: 191-200. http://www.ncbi.nlm.nih.gov/pubmed/15072701.

Bleicher KH, Bohm HJ, Muller K, Alanine AI. (2003). Hit and lead generation: Beyond high-throughput screening. *Nature Reviews Drug Discovery* 2: 369-378. http://www.ncbi.nlm.nih.gov/pubmed/12750740.

Blount BC, Valentin-Blasini L, Osterloh JD, Mauldin JP, Pirkle JL. (2007). Perchlorate exposure of the US Population, 2001-2002. *Journal of Exposure Science and Environmental Epidemiology* 17: 400-407. http://www.ncbi.nlm.nih.gov/pubmed/17051137.

Boekelheide K, Blumberg B, Chapin RE, Cote I, Graziano JH, Janesick A, et al. (2012). Predicting later-life outcomes of early-life exposures. *Environmental Health Perspectives* 120: 1353-1361. http://dx.doi.org/10.1289/ehp.1204934.

Bois FY, Jamei M, Clewell HJ. (2010). PBPK modelling of inter-individual variability in the pharmacokinetics of environmental chemicals. *Toxicology* 278: 256-267. http://www.ncbi.nlm.nih.gov/pubmed/20600548.

Bollati V, Baccarelli A. (2010). Environmental epigenetics. *Heredity (Edinb)* 105: 105-112. http://www.ncbi.nlm.nih.gov/pubmed/20179736.

Boobis AR, Doe JE, Heinrich-Hirsch B, Meek ME, Munn S, Ruchirawat M, et al. (2008). IPCS framework for analyzing the relevance of a noncancer mode of action for humans. *Crit Rev Toxicol* 38: 87-96. http://www.ncbi.nlm.nih.gov/pubmed/18259981.

Borgelt C, Kruse R. (2002). Induction of Association Rules: Apriori Implementation. In *Compstat* (pp. 59): Physica-Verlag HD. http://dx.doi.org/10.1007/978-3-642-57489-4_59.

BOSC (Board of Scientific Counselors). (2010). *Review of the EPA Office of Research and Development's (ORD) Computational Toxicology Research Program (CTRP)*. Washington, DC. Retrieved from http://www.epa.gov/osp/bosc/pdf/ctox1004rpt.pdf (accessed August 29, 2014).

Bosson J, Stenfors N, Bucht A, Helleday R, Pourazar J, Holgate ST, et al. (2003). Ozone-induced bronchial epithelial cytokine expression differs between healthy and asthmatic subjects. *Clinical and Experimental Allergy* 33: 777-782. http://www.ncbi.nlm.nih.gov/pubmed/12801312.

Brown A, Bickley LK, Le Page G, Hosken DJ, Paull GC, Hamilton PB, et al. (2011). Are toxicological responses in laboratory (inbred) zebrafish representative of those in outbred (wild) populations? - A case study with an endocrine disrupting chemical. *Environmental Science and Technology* 45: 4166-4172. http://www.ncbi.nlm.nih.gov/pubmed/21469706.

Bucher JR, Thayer K, Birnbaum LS. (2011). The Office of Health assessment and translation: A problem-solving resource for the National Toxicology Program. *Environmental Health Perspectives* 119: A196-197. http://www.ncbi.nlm.nih.gov/pubmed/21531652.

Burgess-Herbert SL, Euling SY. (2013). Use of comparative genomics approaches to characterize interspecies differences in response to environmental chemicals: Challenges, opportunities, and research needs. *Toxicology and Applied Pharmacology* 271: 372-385. http://www.ncbi.nlm.nih.gov/pubmed/22142766.

Burgoon LD. (2011). Potential of Genomic Data on PAHs to Inform Cumulative Assessment. *National Academy of Sciences Meeting: Mixtures and Cumulative Risk Assessment*. http://nas-sites.org/emergingscience/meetings/mixtures/workshop-presentations-mixtures/.

Burgoon LD, Zacharewski TR. (2008). Automated quantitative dose-response modeling and point of departure determination for large toxicogenomic and high-throughput screening data sets. *Toxicological Sciences* 104: 412-418. http://www.ncbi.nlm.nih.gov/pubmed/18441342.

Bush WS, Moore JH. (2012). Chapter 11: Genome-wide association studies. *Public Library of Science Computational Biology* 8. http://www.ncbi.nlm.nih.gov/pubmed/23300413.

Califano A, Butte AJ, Friend S, Ideker T, Schadt E. (2012). Leveraging models of cell regulation and GWAS data in integrative network-based association studies. *Nature Genetics* 44: 841-847. http://www.ncbi.nlm.nih.gov/pubmed/22836096.

Chadwick LH. (2012). The NIH roadmap epigenomics program data resource. *Epigenomics* 4: 317-324. http://www.ncbi.nlm.nih.gov/pubmed/22690667.

Chandler K, Barrier M, Jeffay S, Nichols HP, Kleinstreuer N. (2011). Evaluation of 309 environmental chemicals using a mouse embryonic stem cell adherent cell differentiation and cytotoxicity assay. *PloS One* 6: e18540. http://www.ncbi.nlm.nih.gov/pubmed/21666745.

Chen J, Blackwell TW, Fermin D, Menon R, Chen Y, Gao J, et al. (2007). Evolutionary-conserved gene expression response profiles across mammalian tissues. *OMICS* 11: 96-115. http://www.ncbi.nlm.nih.gov/pubmed/17411398.

Chen Y, Jin JY, Mukadam S, Malhi V, Kenny JR. (2012). Application of IVIVE and PBPK modeling in prospective prediction of clinical pharmacokinetics: Strategy and approach during the drug discovery phase with four case studies. *Biopharmaceutics & Drug Disposition* 33: 85-98. http://www.ncbi.nlm.nih.gov/pubmed/22228214.

Chiu WA, Euling SY, Scott CS, Subramaniam RP. (2010). Approaches to advancing quantitative human health risk assessment of environmental chemicals in the post-genomic era. *Toxicology and Applied Pharmacology*. http://www.ncbi.nlm.nih.gov/pubmed/20353796.

Chiu WA, Euling SY, Scott CS, Subramaniam RP. (2013). Approaches to advancing quantitative human health risk assessment of environmental chemicals in the post-genomic era. *Toxicol Appl Pharmacol* 271: 309-323. http://www.ncbi.nlm.nih.gov/pubmed/20353796.

Chiu WA, Okino MS, Evans MV. (2009). Characterizing uncertainty and population variability in the toxicokinetics of trichloroethylene and metabolites in mice, rats, and humans using an updated database, physiologically based pharmacokinetic (PBPK) model, and Bayesian approach. *Toxicology and Applied Pharmacology* 241: 36-60. http://www.ncbi.nlm.nih.gov/pubmed/19660485.

Chuang HY, Lee E, Liu YT, Lee D, Ideker T. (2007). Network-based classification of breast cancer metastasis. *Molecular Systems Biology* 3: 140. http://www.ncbi.nlm.nih.gov/pubmed/17940530.

Clewell HJ, 3rd, Andersen ME, Barton HA. (2002). A consistent approach for the application of pharmacokinetic modeling in cancer and noncancer risk assessment. *Environ Health Perspect* 110: 85-93. http://www.ncbi.nlm.nih.gov/pubmed/11781169.

Cohen Hubal EA, Richard A, Aylward L, Edwards S, Gallagher J, Goldsmith MR, et al. (2010). Advancing exposure characterization for chemical evaluation and risk assessment. *Journal*

of Toxicology and Environmental Health Part B: Critical Reviews 13: 299-313. http://www.ncbi.nlm.nih.gov/pubmed/20574904.
Collins M. (2009). Generating 'omic knowledge': The role of informatics in high content screening. *Combinatorial Chemistry & High Throughput Screening* 12: 917-925. http://www.ncbi.nlm.nih.gov/pubmed/19531005.
Cordell HJ. (2009). Detecting gene-gene interactions that underlie human diseases. *Nature Reviews Genetics* 10: 392-404. http://www.ncbi.nlm.nih.gov/pubmed/19434077.
Cortessis VK, Thomas DC, Levine AJ, Breton CV, Mack TM, Siegmund KD, et al. (2012). Environmental epigenetics: Prospects for studying epigenetic mediation of exposure-response relationships. *Human Genetics* 131: 1565-1589. http://www.ncbi.nlm.nih.gov/pubmed/22740325.
Cote I, Anastas PT, Birnbaum LS, Clark RM, Dix DJ, Edwards SW, et al. (2012). Advancing the next generation of health risk assessment. *Environmental Health Perspectives* 120: 1499-1502. http://www.ncbi.nlm.nih.gov/pubmed/22875311.
Crawford-Brown D. (2013). The role of advanced biological data in the rationality of risk-based regulatory decisions. *Journal of Environmental Protection* 4: 238-249. http://www.scirp.org/journal/PaperInformation.aspx?PaperID=28638.
Crofton KM. (2008). Thyroid disrupting chemicals: Mechanisms and mixtures. *Int J Androl* 31: 209-223. http://www.ncbi.nlm.nih.gov/pubmed/18217984.
Crofton KM, Zoeller RT. (2005). Mode of action: Neurotoxicity induced by thyroid hormone disruption during development--hearing loss resulting from exposure to PHAHs. *Critical Reviews in Toxicology* 35: 757-769. http://www.ncbi.nlm.nih.gov/pubmed/16417043.
Cronican AA, Fitz NF, Carter A, Saleem M, Shiva S, Barchowsky A, et al. (2013). Genome-wide alteration of histone H3K9 acetylation pattern in mouse offspring prenatally exposed to arsenic. *Public Library of Science One* 8: e53478. http://www.ncbi.nlm.nih.gov/pubmed/23405071.
Crump KS, Chen C, Louis TA. (2010a). The future use of *in vitro* data in risk assessment to set human exposure standards: Challenging problems and familiar solutions. *Environmental Health Perspectives* 118: 1350-1354. http://www.ncbi.nlm.nih.gov/pubmed/20562051.
Crump KS, Chiu WA, Subramaniam RP. (2010b). Issues in using human variability distributions to estimate low-dose risk. *Environ Health Perspect* 118: 387-393. http://www.ncbi.nlm.nih.gov/pubmed/20064772.
Crump KS, Hoel DG, Langley CH, Peto R. (1976). Fundamental carcinogenic processes and their implications for low dose risk assessment. *Cancer Res* 36: 2973-2979. http://www.ncbi.nlm.nih.gov/pubmed/975067.
De Coster S, van Larebeke N. (2012). Endocrine-disrupting chemicals: Associated disorders and mechanisms of action. *Journal of Environmental Public Health* 2012: 713696. http://www.ncbi.nlm.nih.gov/pubmed/22991565.
Dearry A. (2013). Integrating environmental health data to advance discovery. Presentation. *National Academy of Sciences Meeting: Integrating Environmental Health Data to Advance Discovery* January 10-11. http://www.youtube.com/watch?v=5TIVa2x4x7c&list=PLzsdEyVNFvgyizsegxlcIbLz1glyOIHxJ&index=2.
Degitz SJ, Holcombe GW, Flynn KM, Kosian PA, Korte JJ, Tietge JE. (2005). Progress towards development of an amphibian-based thyroid screening assay using Xenopus laevis. Organismal and thyroidal responses to the model compounds 6-propylthiouracil, methimazole, and thyroxine. *Toxicological Sciences* 87: 353-364. http://www.ncbi.nlm.nih.gov/pubmed/16002479.
Dellarco VL, McGregor D, Berry SC, Cohen SM, Boobis AR. (2006). Thiazopyr and thyroid disruption: Case study within the context of the 2006 IPCS Human Relevance Framework for analysis of

a cancer mode of action. *Crit Rev Toxicol* 36: 793-801. http://www.ncbi.nlm.nih.gov/pubmed/17118729.
Derry JM, Mangravite LM, Suver C, Furia MD, Henderson D, Schildwachter X, et al. (2012). Developing predictive molecular maps of human disease through community-based modeling. *Nature Genetics* 44: 127-130. http://www.ncbi.nlm.nih.gov/pubmed/22281773.
Devlin RB. (2014). Using Ozone to Validate a Systems Biology Approach to Toxicity Testing.HEI Annual Meeting, May 4-6, Alexandria, VA.
Devlin RB, Duncan KE, Jardim M, Schmitt MT, Rappold AG, Diaz-Sanchez D. (2012). Controlled exposure of healthy young volunteers to ozone causes cardiovascular effects. *Circulation* 126: 104-111. http://www.ncbi.nlm.nih.gov/pubmed/22732313.
DeWoskin RS, Knudsen TB, Shah I. (2014). Virtual Models (vM). In *Encyclopedia of Toxicology 3rd edition vol 4*: Academic Press.
DeWoskin RS, Thompson CM. (2008). Renal clearance parameters for PBPK model analysis of early lifestage differences in the disposition of environmental toxicants. *Regul Toxicol Pharmacol* 51: 66-86. http://www.ncbi.nlm.nih.gov/pubmed/18433959.
Diabetes Genetics Initiative of Broad Institute of Harvard, MIT LU, Novartis Institutes of BioMedical Research, Saxena R, Voight BF, Lyssenko V, et al. (2007). Genome-wide association analysis identifies loci for type 2 diabetes and triglyceride levels. *Science* 316: 1331-1336. http://www.sciencemag.org/content/316/5829/1331.abstract.
Diaz Ochoa JG, Bucher J, Pery AR, Zaldivar Comenges JM, Niklas J, Mauch K. (2012). A multi-scale modeling framework for individualized, spatiotemporal prediction of drug effects and toxicological risk. *Front Pharmacol* 3: 204. http://www.ncbi.nlm.nih.gov/pubmed/23346056.
Dick E, Rajamohan D, Ronksley J, Denning C. (2010). Evaluating the utility of cardiomyocytes from human pluripotent stem cells for drug screening. *Biochemical Society Transactions* 38: 1037-1045. http://www.ncbi.nlm.nih.gov/pubmed/20659000.
Dix DJ, Houck KA, Martin MT, Richard AM, Setzer RW, Kavlock RJ. (2007). The ToxCast program for prioritizing toxicity testing of environmental chemicals. *Toxicol Sci* 95: 5-12. http://www.ncbi.nlm.nih.gov/pubmed/16963515.
Duncan KE, Crooks J, Miller DJ, Burgoon L, Schmitt MT, Edwards S, et al. (2013). Temporal profile of gene expression alterations in primary human bronchial epithelial cells following *in vivo* exposure to ozone *American Thoracic Society International Conference: D105 Genetics and Epigenetics of Lung Disease* Meeting Abstract A5866. http://www.atsjournals.org/doi/abs/10.1164/ajrccm-conference.2013.187.1_MeetingAbstracts.A5866.
Duncan KE, Dailey LA, Carson JL, Hernandez ML, Peden DB, Devlin RB. (2012). Cultured basal airway epithelial cells from asthmatics display baseline gene expression profiles that differ from normal healthy cells and exhibit differential responses to ambient air pollution particles *American Journal of Respiratory and Critical Care Medicine* 185: A4291. http://www.atsjournals.org/doi/abs/10.1164/ajrccm-conference.2012.185.1_MeetingAbstracts.A4291.
EC (European Commission). (2010). *Main Findings of the Report: Alternative (Non-animal) Methods for Cosmetics Testing: Current Status and Future Prospects*. Institute for Health and Consumer Protection. Retrieved from http://www.iivs.org/workspace/assets/news-assets/findings_cosmetics_2011.pdf (accessed August 29, 2014).
EC (European Commission). (2011). *4th Report on the Implementation of the "Community Strategy for Endocrine Disruptors" a Range of Substances Suspected of Interfering with the Hormone Systems of Humans and Wildlife*. (COM (1999) 706). Brussels. Retrieved from http://ec.europa.eu/environment/chemicals/endocrine/pdf/sec_2011_1001.pdf (accessed August 29, 2014).

EC (European Commission). (2013). *SEURAT - 1:Towards the replacement of in vivo repeated dose systemic toxicity testing (Volume 3)*. Retrieved from http://www.seurat-1.eu/media/download_gallery/SEURAT%201-v3%28bd.pdf (accessed August 29, 2014).

ECETOC (European Centre for Ecotoxicology and Toxicology of Chemicals). (2013). *Use of category approaches, read-across and (Q)SAR: General considerations.* . Retrieved from http://www.ecetoc.org/index.php?mact=MCSoap,cntnt01,details,0&cntnt01by_category=22&cntnt01order_by=date%20Desc&cntnt01template=display_list_v2&cntnt01display_template=display_details_v2&cntnt01document_id=7423&cntnt01returnid=59 (accessed August 29, 2014).

ECHA (European Chemicals Agency). (2012). *General Report 2011: The Year of Dissemination*. Helsinki. Retrieved from http://echa.europa.eu/documents/10162/13560/mb_06_2012_general_report_2011_final_en.pdf (accessed August 29, 2014).

ECHA (European Chemicals Agency). (2013). *Proposal for identification of a substance as a CMR 1a or 1b, PBT, vPvB or a substance of an equivalent level of concern for 4-Nonylphenol, branched and linear, ethoxylated*. Dortmund. Retrieved from http://echa.europa.eu/documents/10162/13638/annex_xv_svhc_4nonylphenol_en.pdf (accessed August 29, 2014).

ECHA (European Chemicals Agency). (2014). *About Us*. Retrieved from http://echa.europa.eu/about-us (accessed August 29, 2014).

Edwards SW, Preston RJ. (2008). Systems biology and mode of action based risk assessment. *Toxicological Sciences* 106: 312-318. http://www.ncbi.nlm.nih.gov/pubmed/18791183.

Ein-Dor L, Kela I, Getz G, Givol D, Domany E. (2005). Outcome signature genes in breast cancer: Is there a unique set? *Bioinformatics* 21: 171-178. http://www.ncbi.nlm.nih.gov/pubmed/15308542.

Eisenberg MC, Samuels M, DiStefano JJ, 3rd. (2008). Extensions, validation, and clinical applications of a feedback control system simulator of the hypothalamo-pituitary-thyroid axis. *Thyroid* 18: 1071-1085. http://www.ncbi.nlm.nih.gov/pubmed/18844475.

El Muayed M, Raja MR, Zhang X, MacRenaris KW, Bhatt S, Chen X, et al. (2012). Accumulation of cadmium in insulin-producing β cells. *Islets* 4: 405-416. http://www.ncbi.nlm.nih.gov/pubmed/23466887.

Ellinger-Ziegelbauer H, Gmuender H, Bandenburg A, Ahr HJ. (2008). Prediction of a carcinogenic potential of rat hepatocarcinogens using toxicogenomics analysis of short-term *in vivo* studies. *Mutation Research* 637: 23-39. http://dx.doi.org/10.1016/j.mrfmmm.2007.06.010.

Emdin SO, Dodson GG, Cutfield JM, Cutfield SM. (1980). Role of zinc in insulin biosynthesis. Some possible zinc-insulin interactions in the pancreatic B-cell. *Diabetologia* 19: 174-182. http://www.ncbi.nlm.nih.gov/pubmed/6997118.

ENCODE Project Consortium. (2012). An integrated encyclopedia of DNA elements in the human genome. *Nature* 489: 57-74. http://www.ncbi.nlm.nih.gov/pubmed/22955616.

EPA (U.S. Environmental Protection Agency). (1995). *The Use of the Benchmark Dose Approach in Health Risk Assessment*. (EPA/630/R-94/007). Washington, DC. Retrieved from http://www.epa.gov/raf/publications/pdfs/BENCHMARK.PDF (accessed August 29, 2014).

EPA (U.S. Environmental Protection Agency). (2000). *Benzene (CASRN 71-43-2) IRIS Summary. II. Carcinogenicity Assessment for Lifetime Exposure*. Retrieved from http://www.epa.gov/iris/subst/0276.htm (accessed August 29, 2014).

EPA (U.S. Environmental Protection Agency). (2005). *Guidelines for carcinogen risk assessment*. (EPA/630/P-03/001F). Retrieved from http://www.epa.gov/cancerguidelines/ (accessed August 29, 2014).

EPA (U.S. Environmental Protection Agency). (2006). *Approaches for the Application of Physiologically based Pharmacokinetic (PBPK) Models and Supporting Data in Risk*

Assessment. Final Report. (EPA/600/R-05/043A). Retrieved from http://cfpub.epa.gov/ncea/cfm/recordisplay.cfm?deid=157668 (accessed August 29, 2014).

EPA (U.S. Environmental Protection Agency). (2008). *Uncertainty and Variability in Physiologically Based Pharmacokinetic Models: Key Issues and Case Studies*. Washington, DC. Retrieved from http://ofmpub.epa.gov/eims/eimscomm.getfile?p_download_id=477286.

EPA (U.S. Environmental Protection Agency). (2009a). *Integrated Science Assessment for Particulate Matter (Final Report)*. Washington, DC. Retrieved from http://cfpub.epa.gov/ncea/cfm/recordisplay.cfm?deid=216546 (accessed August 29, 2014).

EPA (U.S. Environmental Protection Agency). (2009b). *Strategic Plan for the Future of Toxicity Testing and Risk Assessment at the U.S. EPA*. Washington, DC: Office of the Science Advisor, Science Policy Council. Retrieved from http://www.epa.gov/spc/toxicitytesting/docs/toxtest_strategy_032309.pdf (accessed August 29, 2014).

EPA (U.S. Environmental Protection Agency). (2010). *Advancing the Next Generation (NexGen) of Risk Assessment: The Prototypes Workshop*. Research Triangle Park, NC. Retrieved from http://www.epa.gov/risk/nexgen/docs/NexGen-Prototypes-Workshop-Summary.pdf (accessed August 29, 2014).

EPA (U.S. Environmental Protection Agency). (2011a). *Advancing the Next Generation (NexGen) of Risk Assessment: Public Dialogue Conference Report*. Washington, DC. Retrieved from http://www.epa.gov/risk/nexgen/docs/NexGen-Public-Conf-Summary.pdf (accessed August 29, 2014).

EPA (U.S. Environmental Protection Agency). (2011b). *Framework for an EPA Chemical Safety for Sustainability Research Program*. Washington, DC: Office of Research and Development. Retrieved from http://www.epa.gov/ord/priorities/docs/CSSFramework.pdf (accessed August 29, 2014).

EPA (U.S. Environmental Protection Agency). (2011c). *The Incorporation of In Silico Models and In Vitro High Throughput Assays in the Endocrine Disruptor Screening Program (EDSP) for Prioritization and Screening*. Washington, DC. Retrieved from http://www.epa.gov/endo/pubs/edsp21_work_plan_summary%20_overview_final.pdf (accessed August 29, 2014).

EPA (U.S. Environmental Protection Agency). (2011d). *Use of "Omic" Technology to Inform the Risk Assessment, Support Document for Case Study: Propiconazole. Appendices A and B*. Washington, DC: Federal Insecticide, Fungicide, and Rodenticide Act Scientific Advisory Panel (FIFRA SAP). Retrieved from http://www.regulations.gov/#!documentDetail;D=EPA-HQ-OPP-2011-0284-0003 (accessed August 29, 2014).

EPA (U.S. Environmental Protection Agency). (2012a). *Benchmark Dose Technical Guidance*. (EPA/100/R-12/001). Retrieved from http://www.epa.gov/raf/publications/pdfs/benchmark_dose_guidance.pdf (accessed August 29, 2014).

EPA (U.S. Environmental Protection Agency). (2012b). *Chemical Safety for Sustainability Strategic Research Action Plan 2012-2016*. Office of Research and Development. Retrieved from http://epa.gov/research/docs/css-strap.pdf (accessed August 29, 2014).

EPA (U.S. Environmental Protection Agency). (2012c). *Computational Toxicology Research: Overview Materials for EPA Science Advisory Board Exposure & Human Health Committee*. Washington, DC. Retrieved from http://yosemite.epa.gov/sab/sabproduct.nsf/AEACDCDADEDFAAF3852579F2005A4C5A/$File/EPA+CompTox+SAB+Materials+4-26-2012+v4.pdf (accessed August 29, 2014).

EPA (U.S. Environmental Protection Agency). (2012d). *Human Health Risk Assessment Strategic Research Action Plan 2012-2016*. Office of Research and Development. Retrieved from http://epa.gov/research/docs/hhra-strap.pdf (accessed August 29, 2014).
EPA (U.S. Environmental Protection Agency). (2013a). *Advancing the Next Generation (NexGen) of Risk Assessment*. Retrieved from http://www.epa.gov/risk/nexgen/ (accessed August 29, 2014).
EPA (U.S. Environmental Protection Agency). (2013b). *Basic Information About Benzene in Drinking Water*. Retrieved from http://water.epa.gov/drink/contaminants/basicinformation/benzene.cfm (accessed August 29, 2014).
EPA (U.S. Environmental Protection Agency). (2013c). *Integrated Science Assessment of Ozone and Related Photochemical Oxidants (Final Report)*. Washington, DC. Retrieved from http://cfpub.epa.gov/ncea/isa/recordisplay.cfm?deid=247492 (accessed March 28, 2013).
EPA (U.S. Environmental Protection Agency). (2013d). *IRIS Toxicological review of benzo[a]pyrene (Public comment draft)*. Retrieved from http://cfpub.epa.gov/ncea/iris_drafts/recordisplay.cfm?deid=66193 (accessed August 29, 2014).
EPA (U.S. Environmental Protection Agency). (2013e). *Lead Integrated Science Assessment*. (EPA/600/R-10/075F). Retrieved from http://cfpub.epa.gov/ncea/cfm/recordisplay.cfm?deid=255721.
EPA (U.S. Environmental Protection Agency). (2013f). *Sustainable futures: Models & methods*. Retrieved from http://www.epa.gov/oppt/sf/tools/methods.htm (accessed August 29, 2014).
EPA (U.S. Environmental Protection Agency). (2014a). *About the Chemical Safety for Sustainability Research Program*. Retrieved from http://www2.epa.gov/aboutepa/about-chemical-safety-sustainability-research-program (accessed August 29. 2014).
EPA (U.S. Environmental Protection Agency). (2014b). *About the Human Health Risk Assessment Program*. Retrieved from http://www2.epa.gov/aboutepa/about-human-health-risk-assessment-program (accessed August 29, 2014).
EPA (U.S. Environmental Protection Agency). (2014c). *About the Office of Research and Development (ORD)*. Retrieved from http://www2.epa.gov/aboutepa/about-office-research-and-development-ord (accessed August 29, 2014).
EPA (U.S. Environmental Protection Agency). (2014d). *ACToR*. Retrieved from http://actor.epa.gov/actor/faces/ACToRHome.jsp (accessed August 29, 2014).
EPA (U.S. Environmental Protection Agency). (2014e). *Advancing the Next Generation (NexGen) of Risk Assessment: Conferences and Workshops*. Washington, DC. Retrieved from http://www.epa.gov/risk/nexgen/workshops.htm (accessed August 29, 2014).
EPA (U.S. Environmental Protection Agency). (2014f). *Basic information about the Toxic Substance Control Act Inventory*. Retrieved from http://www.epa.gov/oppt/existingchemicals/pubs/tscainventory/basic.html#background (accessed August 29, 2014).
EPA (U.S. Environmental Protection Agency). (2014g). *Computational Toxicology Research*. Retrieved from http://www.epa.gov/ncct/ (accessed August 29, 2014).
EPA (U.S. Environmental Protection Agency). (2014h). *Endocrine Disruptor Screening Program (EDSP)*. Retrieved from http://www.epa.gov/endo/ (accessed August 29, 2014).
EPA (U.S. Environmental Protection Agency). (2014i). *Framework for Human Health Risk Assessment to Inform Decision Making*. (EPA-HQ-ORD-2012-0579.). Office of the Science Advisor. Retrieved from http://www.epa.gov/raf/frameworkhhra.htm.
EPA (U.S. Environmental Protection Agency). (2014j). *ToxCast Data*. Retrieved from http://www.epa.gov/ncct/toxcast/data.html (accessed August 29, 2014).

EPA (U.S. Environmental Protection Agency). (2014k). *ToxCastDB Data Collection List: In Vitro Assays*. Retrieved from http://actor.epa.gov/actor/faces/ToxCastDB/DataCollectionList.jsp (accessed August 29, 2014).
EPA (U.S. Environmental Protection Agency). (2014l). *ToxCast™*. Retrieved from http://www.epa.gov/ncct/toxcast/ (accessed August 29, 2014).
EPA (U.S. Environmental Protection Agency). (2014m). *The Virtual Embryo Project (v-Embryo™)*. Retrieved from http://www.epa.gov/ncct/v-Embryo/index.html (accessed August 29, 2014).
EPA (U.S. Environmental Protection Agency). (2014n). *The Virtual Liver Project (v-Liver™)*. Retrieved from http://www.epa.gov/ncct/virtual_liver/ (accessed August 29, 2014).
EPA (U.S. Environmental Protection Agency). (2014o). *Vocabulary Catalog List Detail - Integrated Risk Information System (IRIS) Glossary* Retrieved from http://ofmpub.epa.gov/sor_internet/registry/termreg/searchandretrieve/glossariesandkeywordlists/search.do?details=&glossaryName=IRIS%20Glossary (accessed August 29, 2014).
Eskenazi B, Bradman A, Gladstone EA, Jaramillo S, Birch K, Holland N. (2003). CHAMACOS, A longitudinal birth cohort study: Lessons from the fields. *Journal of Children's Health* 1: 3-27. http://ehs.sph.berkeley.edu/holland/publications/files/Eskenazi2003.pdf
Fielden MR, Adai A, Dunn RT, Olaharski A, Searfoss G, Sina J, et al. (2011). Development and evaluation of a genomic signature for the prediction and mechanistic assessment of nongenotoxic hepatocarcinogens in the rat. *Toxicological Sciences* 124: 54-74. http://dx.doi.org/10.1093/toxsci/kfr202.
Fielden MR, Brennan R, Gollub J. (2007). A gene expression biomarker provides early prediction and mechanistic assessment of hepatic tumor induction by nongenotoxic chemicals. *Toxicological Sciences* 99: 90-100. http://dx.doi.org/10.1093/toxsci/kfm156.
Fielden MR, Eynon BP, Natsoulis G, Jarnagin K, Banas D, Kolaja KL. (2005). A gene expression signature that predicts the future onset of drug-induced renal tubular toxicity. *Toxicologic Pathology* 33: 675-683. http://dx.doi.org/10.1080/01926230500321213.
Fielden MR, Nie A, McMillian M, Elangbam CS, Trela BA, Yang Y, et al. (2008). Interlaboratory evaluation of genomic signatures for predicting carcinogenicity in the rat. *Toxicological Sciences* 103: 28-34. http://dx.doi.org/10.1093/toxsci/kfn022.
Fini JB, Le Mevel S, Turque N, Palmier K, Zalko D, Cravedi JP, et al. (2007). An *in vivo* multiwell-based fluorescent screen for monitoring vertebrate thyroid hormone disruption. *Environ Sci Technol* 41: 5908-5914. http://www.ncbi.nlm.nih.gov/pubmed/17874805.
Fortunel NO, Otu HH, Ng HH, Chen J, Mu X, Chevassut T, et al. (2003). Comment on "'Stemness': Transcriptional profiling of embryonic and adult stem cells" and "a stem cell molecular signature". *Science* 302: 393; author reply 393. http://www.ncbi.nlm.nih.gov/pubmed/14563990.
Fowler BA, Whittaker MH, Lipsky M, Wang G, Chen XQ. (2004). Oxidative stress induced by lead, cadmium and arsenic mixtures: 30-day, 90-day, and 180-day drinking water studies in rats: An overview. *Biometals* 17: 567-568. http://www.ncbi.nlm.nih.gov/pubmed/15688865.
Frey C, Crawford-Brown D, Zheng J, Loughlin D USE OAQPS. (2003). *Hierarchy of methods to characterize uncertainty: State of science of methods for describing and quantifying uncertainty*. RTP, NC. Retrieved from http://www4.ncsu.edu/~frey/reports/Frey_CrawfordBrown_Zheng_Loughlin_2003.pdf.
Friend S. (2013). Scientific opportunities from heterogeneous biological data analysis: Overcoming complexity. Presentation. *National Academy of Sciences Meeting: Integrating Environmental Health Data to Advance Discovery* January 10-11. http://nas-sites.org/emergingscience/files/2013/01/FRIEND-Jan10-final.pdf.

Froehlicher M, Liedtke A, Groh KJ, Neuhauss SC, Segner H, Eggen RI. (2009). Zebrafish (Danio rerio) neuromast: Promising biological endpoint linking developmental and toxicological studies. *Aquatic Toxicology* 95: 307-319. http://www.ncbi.nlm.nih.gov/pubmed/19467721.

Functional Genomics Data Society. (2014). *Minimum Information About a Microarray Experiment*. Retrieved from http://fged.org/projects/miame/ (accessed August 29, 2014).

Gangwal S, Reif DM, Mosher S, Egeghy PP, Wambaugh JF, Judson RS, et al. (2012). Incorporating exposure information into the toxicological prioritization index decision support framework. *Science of the Total Environment* 435-436: 316-325. http://www.ncbi.nlm.nih.gov/pubmed/22863807.

Garte S, Taioli E, Popov T, Bolognesi C, Farmer P, Merlo F. (2008). Genetic susceptibility to benzene toxicity in humans. *Journal of Toxicology and Environmental Health, Part A* 71: 1482-1489. http://www.ncbi.nlm.nih.gov/pubmed/18836923.

Gaylor DW, Kodell RL, Chen JJ, Krewski D. (1999). A unified approach to risk assessment for cancer and noncancer endpoints based on benchmark doses and uncertainty/safety factors. *Regul Toxicol Pharmacol* 29: 151-157. http://www.ncbi.nlm.nih.gov/pubmed/10341145.

George BJ, Schultz BD, Palma T, Vette AF, Whitaker DA, Williams RW. (2011). *An Evaluation of EPA's National-Scale Air Toxics Assessment (NATA): Comparison with Benzene Measurements in Detroit, Michigan*. U.S. Environmental Protection Agency. Retrieved from http://cfpub.epa.gov/si/si_public_record_report.cfm?dirEntryId=221831&fed_org_id=770&SIType=PR&TIMSType=&showCriteria=0&address=nerl/pubs.html&view=citation&personID=6172&role=Author&sortBy=pubDateYear&count=100&dateBeginPublishedPresented= (accessed August 29, 2014).

German Federal Ministry for Education and Research. (2014). *Virtual Liver Network*. Retrieved from http://www.virtual-liver.de/wordpress/en (accessed August 29, 2014).

Gibbs-Flournoy EA, Simmons SO, Bromberg PA, Dick TP, Samet JM. (2013). Monitoring intracellular redox changes in ozone-exposed airway epithelial cells. *Environmental Health Perspectives* 121: 312-317. http://www.ncbi.nlm.nih.gov/pubmed/23249900.

Gibson GG, Rostami-Hodjegan A. (2007). Modelling and simulation in prediction of human xenobiotic absorption, distribution, metabolism and excretion (ADME): *In vitro-in vivo* extrapolations (IVIVE). *Xenobiotica* 37: 1013-1014. http://www.ncbi.nlm.nih.gov/pubmed/17968734.

Gieseck RL, Hannan NR, Bort R, Hanley NA, Drake RA, Cameron GW, et al. (2014). Maturation of induced pluripotent stem cell derived hepatocytes by 3D-culture. *PloS One* 9: 22. http://www.ncbi.nlm.nih.gov/pubmed/24466060.

Godderis L, Thomas R, Hubbard AE, Tabish AM, Hoet P, Zhang L, et al. (2012). Effect of chemical mutagens and carcinogens on gene expression profiles in human TK6 cells. *Public Library of Science One* 7: e39205. http://europepmc.org/abstract/MED/22723965.

Golbraikh A, Wang XS, Zhu H, Tropsha A. (2012). Predictive QSAR modeling: Methods and applications in drug discovery and chemical risk assessment. *Handbook of Computational Chemistry*1309-1342. http://link.springer.com/referenceworkentry/10.1007%2F978-94-007-0711-5_37#.

Gold LS, Manley NB, Slone TH, Ward JM. (2001). Compendium of chemical carcinogens by target organ: Results of chronic bioassays in rats, mice, hamsters, dogs, and monkeys. *Toxicologic Pathology* 29: 639-652. http://www.ncbi.nlm.nih.gov/pubmed/11794380.

Goldsmith MR, Grulke CM, Brooks RD, Transue TR, Tan YM, Frame A, et al. (2014). Development of a consumer product ingredient database for chemical exposure screening and prioritization. *Food Chem Toxicol* 65: 269-279. http://www.ncbi.nlm.nih.gov/pubmed/24374094.

Goldsmith MR, Peterson SD, Chang DT, Transue TR, Tornero-Velez R, Tan YM, et al. (2012). Informing mechanistic toxicology with computational molecular models. *Methods in Molecular Biology* 929: 139-165. http://www.ncbi.nlm.nih.gov/pubmed/23007429.

Goldstein BD. (1988). Benzene toxicity. *Occupational Medicine* 3: 541-554. http://www.ncbi.nlm.nih.gov/pubmed/3043738.
Goodman JE, Boyce CP, Pizzurro DM, Rhomberg LR. (2014). Strengthening the foundation of next generation risk assessment. *Regul Toxicol Pharmacol* 68: 160-170. http://www.ncbi.nlm.nih.gov/pubmed/24342233.
Greenawalt DM, Sieberts SK, Cornelis MC, Girman CJ, Zhong H, Yang X, et al. (2012). Integrating genetic association, genetics of gene expression, and single nucleotide polymorphism set analysis to identify susceptibility Loci for type 2 diabetes mellitus. *American Journal of Epidemiology* 176: 423-430. http://www.ncbi.nlm.nih.gov/pubmed/22865700.
Grulke CM, Holm K, Goldsmith MR, Tan YM. (2013). PROcEED: Probabilistic reverse dosimetry approaches for estimating exposure distributions. *Bioinformation* 9: 707-709. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3732445/.
Guryev V, Koudijs MJ, Berezikov E, Johnson SL, Plasterk RH, van Eeden FJ, et al. (2006). Genetic variation in the zebrafish. *Genome Research* 16: 491-497. http://www.ncbi.nlm.nih.gov/pubmed/16533913.
Hahsler M, Grün B, Hornik K. (2005). arules - A computational environment for mining association rules and frequent item sets. *Journal of Statistical Software* 14: 1-25. http://www.jstatsoft.org/v14/i15/paper.
Hanahan D, Weinberg RA. (2000). The hallmarks of cancer. *Cell* 100: 57-70. http://www.ncbi.nlm.nih.gov/pubmed/10647931.
Hanahan D, Weinberg RA. (2011). Hallmarks of cancer: The next generation. *Cell* 144: 646-674. http://www.sciencedirect.com/science/article/pii/S0092867411001279.
Hansen KD, Timp W, Bravo HC, Sabunciyan S, Langmead B, McDonald OG, et al. (2011). Increased methylation variation in epigenetic domains across cancer types. *Nature Genetics* 43: 768-775. http://www.ncbi.nlm.nih.gov/pubmed/21706001.
Harrill AH, Desmet KD, Wolf KK, Bridges AS, Eaddy JS, Kurtz CL, et al. (2012). A mouse diversity panel approach reveals the potential for clinical kidney injury due to DB289 not predicted by classical rodent models. *Toxicological Sciences* 130: 416-426. http://www.ncbi.nlm.nih.gov/pubmed/22940726.
Harrill AH, Ross PK, Gatti DM, Threadgill DW, Rusyn I. (2009). Population-based discovery of toxicogenomics biomarkers for hepatotoxicity using a laboratory strain diversity panel. *Toxicological Sciences* 110: 235-243. http://www.ncbi.nlm.nih.gov/pubmed/19420014.
Hartung T, McBride M. (2011). Food for Thought ... on mapping the human toxome. *ALTEX* 28: 83-93. http://www.ncbi.nlm.nih.gov/pubmed/21625825.
Harwood MD, Neuhoff S, Carlson GL, Warhurst G, Rostami-Hodjegan A. (2013). Absolute abundance and function of intestinal drug transporters: A prerequisite for fully mechanistic *in vitro-in vivo* extrapolation of oral drug absorption. *Biopharm Drug Dispos* 34: 2-28. http://www.ncbi.nlm.nih.gov/pubmed/22927116.
Hatch GE, Duncan KE, Diaz-Sanchez D, Schmitt MT, Ghio AJ, Carraway MS, et al. (in press). Progress in assessing air pollutant risks from *in vitro* exposures: Matching ozone dose and effect in human airway cells. *Toxicol Sci*. [Epub ahead of print]. http://www.ncbi.nlm.nih.gov/pubmed/24928893.
Hatch GE, Slade R, Harris LP, McDonnell WF, Devlin RB, Koren HS, et al. (1994). Ozone dose and effect in humans and rats. A comparison using oxygen-18 labeling and bronchoalveolar lavage. *American Journal of Respiratory and Critical Care Medicine* 150: 676-683. http://www.ncbi.nlm.nih.gov/pubmed/8087337.
Hatzimichael E, Crook T. (2013). Cancer epigenetics: New therapies and new challenges. *Journal of Drug Delivery* 2013: 529312. http://www.ncbi.nlm.nih.gov/pubmed/23533770.

Hester RL, Brown AJ, Husband L, Illiescu R, Pruett D, Summers R, et al. (2011). HumMod: A modeling environment for the simulation of integrative human physiology. *Frontiers in Physiology* 2. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3082131/.

Hill AB. (1965). The environment and disease: Association or causation? *Proc R Soc Med* 58: 295-300. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1898525/.

Hines RN. (2007). Ontogeny of human hepatic cytochromes P450. *J Biochem Mol Toxicol* 21: 169-175. http://www.ncbi.nlm.nih.gov/pubmed/17936930.

Hines RN. (2013). Developmental expression of drug metabolizing enzymes: Impact on disposition in neonates and young children. *Int J Pharm* 452: 3-7. http://www.ncbi.nlm.nih.gov/pubmed/22766445.

Holzhutter HG, Drasdo D, Preusser T, Lippert J, Henney AM. (2012). The virtual liver: A multidisciplinary, multilevel challenge for systems biology. *Wiley Interdisciplinary Reviews: Systems Biology and Medicine* 4: 221-235. http://www.ncbi.nlm.nih.gov/pubmed/22246674.

Houck KA, Kavlock RJ. (2008). Understanding mechanisms of toxicity: Insights from drug discovery research. *Toxicology and Applied Pharmacology* 227: 163-178. http://www.ncbi.nlm.nih.gov/pubmed/18063003.

Howlader N, Noone AM, Krapcho M, Garshell J, Neyman N, Altekruse SF, et al. (2013). *SEER Cancer Statistics Review, 1975-2010*. Bethesda, MD: U.S. National Institutes of Health, National Cancer Institute. Retrieved from http://seer.cancer.gov/csr/1975_2010/, based on November 2012 SEER data submission, posted to the SEER web site, April 2013.

Hubal EA. (2009). Biologically relevant exposure science for 21st century toxicity testing. *Toxicological Sciences* 111: 226-232. http://www.ncbi.nlm.nih.gov/pubmed/19602574.

Hunter P, Coveney PV, de Bono B, Diaz V, Fenner J, Frangi AF, et al. (2010). A vision and strategy for the virtual physiological human in 2010 and beyond. *Philosophical Transactions Series A: Mathematical, Physical, and Engineering Sciences* 368: 2595-2614. http://www.ncbi.nlm.nih.gov/pubmed/20439264.

Hunter P, Robbins P, Noble D. (2002). The IUPS human Physiome Project. *Pflugers Archiv* 445: 1-9. http://www.ncbi.nlm.nih.gov/pubmed/12397380.

Hurley PM. (1998). Mode of carcinogenic action of pesticides inducing thyroid follicular cell tumors in rodents. *Environmental Health Perspectives* 106: 437-445. http://www.ncbi.nlm.nih.gov/pubmed/9681970.

IARC (International Agency for Research on Cancer). (2010). *Some Non-heterocyclic Polycyclic Aromatic Hydrocarbons and Some Related Exposures*. Retrieved from http://monographs.iarc.fr/ENG/Monographs/vol92/ (accessed March 25, 2013).

IARC (International Agency for Research on Cancer). (2012). *Monograph 100F: Benzene*. Retrieved from http://monographs.iarc.fr/ENG/Monographs/vol100F/mono100F-24.pdf (accessed August 29, 2014).

Ideker T, Krogan NJ. (2012). Differential network biology. *Mol Syst Biol* 8: 565.

Ilhan G, Karakus S, Andic N. (2006). Risk factors and primary prevention of acute leukemia. *Asian Pacific Journal of Cancer Prevention* 7: 515-517. http://www.ncbi.nlm.nih.gov/pubmed/17250419.

Ingelman-Sundberg M. (2005). Genetic polymorphisms of cytochrome P450 2D6 (CYP2D6): Clinical consequences, evolutionary aspects and functional diversity. *Pharmacogenomics* 5: 6-13. http://www.ncbi.nlm.nih.gov/pubmed/15492763.

Institute of Medicine. (2010). Evaluation of biomarkers and surrogate endpoints in chronic disease. In CM Micheel, JR Ball (Eds.). Washington, DC: The National Academies Press. Retrieved from http://www.nap.edu/openbook.php?record_id=12869.

Irons RD, Chen Y, Wang X, Ryder J, Kerzic PJ. (2013). Acute myeloid leukemia following exposure to benzene more closely resembles de novo than therapy related-disease. *Genes Chromosomes Cancer* 52: 887-894. http://www.ncbi.nlm.nih.gov/pubmed/23840003.

Jack J, Wambaugh JF, Shah I. (2011). Simulating quantitative cellular responses using asynchronous threshold Boolean network ensembles. *BMC Systems Biology* 5: 109. http://www.ncbi.nlm.nih.gov/pubmed/21745399.

Jacobs J. (2011). *Agencies hope robot can speed toxics evaluations, end animal testing*. New York: The New York Times. Retrieved from http://www.nytimes.com/gwire/2011/05/13/13greenwire-agencies-hope-robot-can-speed-toxics-evaluatio-92625.html (accessed August 29, 2014).

Jaspers I, Flescher E, Chen LC. (1997). Ozone-induced IL-8 expression and transcription factor binding in respiratory epithelial cells. *Am J Physiol* 272: L504-511. http://www.ncbi.nlm.nih.gov/pubmed/9124608.

Jiang X, Kumar K, Hu X, Wallqvist A, Reifman J. (2008). DOVIS 2.0: An efficient and easy to use parallel virtual screening tool based on AutoDock 4.0. *Chemical Central Journal* 2: 18. http://www.ncbi.nlm.nih.gov/pubmed/18778471.

JRC (Joint Research Centre). (2014). *Predictive toxicology*. Retrieved from http://ihcp.jrc.ec.europa.eu/our_labs/predictive_toxicology/predictive-toxicology/?searchterm=None (accessed August 29, 2014).

Jubeaux G, Audouard-Combe F, Simon R, Tutundjian R, Salvador A, Geffard O, et al. (2012). Vitellogenin-like proteins among invertebrate species diversity: Potential of proteomic mass spectrometry for biomarker development. *Environmental Science and Technology* 46: 6315-6323. http://www.ncbi.nlm.nih.gov/pubmed/22578134.

Judson R. (2010). *In vitro* screening of environmental chemicals for targeted testing prioritization: The ToxCast project. *Environmental Health Perspectives* 118: 485. http://www.ncbi.nlm.nih.gov/pubmed/20368123.

Judson R, Elloumi F, Setzer W, Li Z, Shah I. (2008). A comparison of machine learning algorithms for chemical toxicity classification using a simulated multi-scale data model. *BMC Bioinformatics* 9: 1-16. http://dx.doi.org/10.1186/1471-2105-9-241.

Judson R, Houck K, Martin M, Knudsen TB, Thomas RS, Sipes N, et al. (2014). *In vitro* and modelling approaches to risk assessment from the U.S. Environmental Protection Agency ToxCast Programme. *Basic & Clinical Pharmacology & Toxicology*n/a-n/a. http://dx.doi.org/10.1111/bcpt.12239.

Judson R, Kavlock R, Martin M, Reif D, Houck K, Knudsen T, et al. (2013). Perspectives on validation of high-throughput assays supporting 21st century toxicity testing. *Alternatives to Animal Experiments* 30: 51-56. http://www.ncbi.nlm.nih.gov/pubmed/23338806.

Judson R, Kavlock RJ, Setzer RW, Hubal EA, Martin MT, Knudsen TB, et al. (2011). Estimating toxicity-related biological pathway altering doses for high-throughput chemical risk assessment. *Chemical Research in Toxicology* 24: 451-462. http://dx.doi.org/10.1021/tx100428e.

Judson R, Martin MT, Egeghy P, Gangwal S, Reif DM, Kothiya P, et al. (2012). Aggregating data for computational toxicology applications: The U.S. Environmental Protection Agency (EPA) Aggregated Computational Toxicology Resource (ACToR) system. *International Journal of Molecular Science* 13: 1805-1831. http://www.ncbi.nlm.nih.gov/pubmed/22408426.

Judson R, Martin MT, Reif DM, Houck KA, Knudsen TB, Rotroff DM, et al. (2010). Analysis of eight oil spill dispersants using rapid, *in vitro* tests for endocrine and other biological activity. *Environmental Science and Technology* 44: 5979-5985. http://dx.doi.org/10.1021/es102150z.

Kanehisa Laboratories. (2014a). *KEGG: Kyoto encyclopedia of genes and genomes*. Retrieved from http://www.genome.jp/kegg/ (accessed August 29, 2014).

Kanehisa Laboratories. (2014b). *KEGG: Kyoto encyclopedia of genes and genomes: Type II Diabetes Mellitus*. Retrieved from http://www.ncbi.nlm.nih.gov/biosystems/83094#show=citations (accessed August 29, 2014).
Kanungo J, Cuevas E, Ali SF, Paule MG. (2014). Zebrafish model in drug safety assessment. *Curr Pharm Des*. http://www.ncbi.nlm.nih.gov/pubmed/24502596.
Kavlock R, Chandler K, Houck K, Hunter S, Judson R, Kleinstreuer N, et al. (2012). Update on EPA's ToxCast program: Providing high throughput decision support tools for chemical risk management. *Chemical Research in Toxicology* 25: 1287-1302. http://www.ncbi.nlm.nih.gov/pubmed/22519603.
Kenyon E. (2012). Interspecies Extrapolation. In *Computational Toxicology* (pp. 19): Humana Press. http://dx.doi.org/10.1007/978-1-62703-050-2_19.
Kleinberg S, Hripcsak G. (2011). A review of causal inference for biomedical informatics. *J Biomed Inform* 44: 1102-1112. http://www.ncbi.nlm.nih.gov/pubmed/21782035.
Kleinstreuer N, Dix D, Rountree M, Baker N, Sipes N, Reif D, et al. (2013a). A computational model predicting disruption of blood vessel development. *PLoS Comput Biol* 9: e1002996. http://www.ncbi.nlm.nih.gov/pubmed/23592958.
Kleinstreuer N, Dix DJ, Houck K, Kavlock R, Knudsen T, Martin M, et al. (2013b). *In vitro* perturbations of targets in cancer hallmark processes predict rodent chemical carcinogenesis. *Environmental Health Perspectives* 131: 40-55. http://toxsci.oxfordjournals.org/content/131/1/40.
Kleinstreuer N, Judson RS, Reif DM, Sipes NS, Singh AV, Chandler KJ, et al. (2011a). Environmental Impact on Vascular Development Predicted by High Throughput Screening. *Environmental Health Perspectives*. http://www.ncbi.nlm.nih.gov/pubmed/21788198.
Kleinstreuer N, Smith AM, West PR, Conard KR, Fontaine BR, Weir-Hauptman AM, et al. (2011b). Identifying developmental toxicity pathways for a subset of ToxCast chemicals using human embryonic stem cells and metabolomics. *Toxicol Appl Pharmacol* 257: 111-121. http://www.ncbi.nlm.nih.gov/pubmed/21925528.
Knudsen T, Daston GP. (2010). Virtual tissues and developmental systems biology (Chapter 23). In T Knudsen, G Daston (Eds.), *Second Edition Comprehensive Toxicology* (pp. 347-358). Oxford, UK: Elsevier Ltd.
Knudsen T, DeWoskin RS. (2011). Systems modeling in development toxicity. In *Handbook of Systems Toxicology*: John Wiley & Sons, Ltd.
Knudsen T, Houck KA, Sipes NS, Singh AV, Judson RS, Martin MT, et al. (2011a). Activity profiles of 309 ToxCast chemicals evaluated across 292 biochemical targets. *Toxicology* 282: 1-15. http://www.ncbi.nlm.nih.gov/pubmed/21251949.
Knudsen T, Kavlock RJ, Daston GP, Stedman D, Hixon M, Kim JH. (2011b). Developmental toxicity testing for safety assessment: New approaches and technologies. *Birth Defects Research: Part B, Developmental and Reproductive Toxicology* 92: 413-420. http://www.ncbi.nlm.nih.gov/pubmed/21770025.
Knudsen T, Kleinstreuer NC. (2011). Disruption of embryonic vascular development in predictive toxicology. *Birth Defects Research: Part C, Embryo Today* 93: 312-323. http://www.ncbi.nlm.nih.gov/pubmed/22271680.
Knudsen T, Martin MT, Kavlock RJ, Judson RJ, Dix DJS, A. V. (2009). Profiling the activity of environmental chemicals in prenatal developmental toxicity studies using the U.S. EPA's ToxRefDB. *Reproductive Toxicology* 28: 209-219. http://www.ncbi.nlm.nih.gov/pubmed/19446433.
Kondraganti SR, Fernandez-Salguero P, Gonzalez FJ, Ramos KS, Jiang W, Moorthy B. (2003). Polycyclic aromatic hydrocarbon-inducible DNA adducts: Evidence by 32P-postlabeling and use of knockout mice for Ah receptor-independent mechanisms of metabolic activation *in*

vivo. *International Journal of Cancer* 103: 5-11. http://www.ncbi.nlm.nih.gov/pubmed/12455047.
Korte JJ, Sternberg RM, Serrano JA, Thoemke KR, Moen SM, Lillegard KE, et al. (2011). Thyroid-stimulating hormone (TSH): Measurement of intracellular, secreted, and circulating hormone in Xenopus laevis and Xenopus tropicalis. *Gen Comp Endocrinol* 171: 319-325. http://www.ncbi.nlm.nih.gov/pubmed/21354158.
Koturbash I, Beland FA, Pogribny IP. (2011). Role of epigenetic events in chemical carcinogenesis--a justification for incorporating epigenetic evaluations in cancer risk assessment. *Toxicology Mechanisms and Methods* 21: 289-297. http://www.ncbi.nlm.nih.gov/pubmed/21495867.
Krewski D, Hogan V, Turner MC, Zeman PL, McDowell I, Edwards N, et al. (2007). An integrated framework for risk management and population health. *Human and Ecological Risk Assessment* 13: 1288-1312. http://www.tandfonline.com/doi/abs/10.1080/10807030701655798?journalCode=bher20#preview.
Krewski D, Westphal M, Al-Zoughool M, Croteau MC, Andersen ME. (2011). New directions in toxicity testing. *Annu Rev Public Health* 32: 161-178.
Krewski D, Westphal M, Andersen ME, Paoli GM, Chiu WA, Al-Zoughool M, et al. (2014). A framework for the next generation of risk science. *Environ Health Perspect* 122: 796-805. http://www.ncbi.nlm.nih.gov/pubmed/24727499.
Lalone CA, Villeneuve DL, Burgoon LD, Russom CL, Helgen HW, Berninger JP, et al. (2013). Molecular target sequence similarity as a basis for species extrapolation to assess the ecological risk of chemicals with known modes of action. *Aquat Toxicol* 144-145: 141-154. http://www.ncbi.nlm.nih.gov/pubmed/24177217.
Lan Q, Zhang L, Li G, Vermeulen R, Weinberg RS, Dosemeci M, et al. (2004). Hematotoxicity in workers exposed to low levels of benzene. *Science* 306: 1774-1776. http://www.ncbi.nlm.nih.gov/pubmed/15576619.
Lau FK, Decastro BR, Mills-Herring L, Tao L, Valentin-Blasini L, Alwis KU, et al. (2013). Urinary perchlorate as a measure of dietary and drinking water exposure in a representative sample of the United States population 2001-2008. *Journal of Exposure Science and Environmental Epidemiology* 23: 207-214. http://www.ncbi.nlm.nih.gov/pubmed/23188482.
Lock EF, Abdo N, Huang R, Xia M, Kosyk O, O'Shea SH, et al. (2012). Quantitative high-throughput screening for chemical toxicity in a population-based *in vitro* model. *Toxicological Sciences* 126: 578-588. http://dx.doi.org/10.1093/toxsci/kfs023.
Lossos IS, Czerwinski DK, Alizadeh AA, Wechser MA, Tibshirani R, Botstein D, et al. (2004). Prediction of survival in diffuse large-B-cell lymphoma based on the expression of six genes. *New England Journal of Medicine* 350: 1828-1837. http://www.ncbi.nlm.nih.gov/pubmed/15115829.
Lutz WK. (1990). Dose-response relationship and low dose extrapolation in chemical carcinogenesis. *Carcinogenesis* 11: 1243-1247. http://www.ncbi.nlm.nih.gov/pubmed/2387009.
Lvovs D, Favorova OO, Favorov AV. (2012). A polygenic approach to the study of polygenic diseases. *Acta Naturae* 4: 59-71. http://www.ncbi.nlm.nih.gov/pubmed/23150804.
Makris SL, Kim JH, Ellis A, Faber W, Harrouk W, Lewis JM, et al. (2011). Current and future needs for developmental toxicity testing. *Birth Defects Research: Part B, Developmental and Reproductive Toxicology* 92: 384-394. http://www.ncbi.nlm.nih.gov/pubmed/21922641.
Martin MT, Judson RS, Reif DM, Kavlock RJ, Dix DJ. (2009a). Profiling chemicals based on chronic toxicity results from the U.S. EPA ToxRef Database. *Environ Health Perspect* 117: 392-399. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2661909/.

Martin MT, Knudsen TB, Judson RS, Kavlock RJ, Dix DJ. (2012). Economic benefits of using adaptive predictive models of reproductive toxicity in the context of a tiered testing program. *Syst Biol Reprod Med* 58: 3-9.
Martin MT, Knudsen TB, Reif DM, Houck KA, Judson RS, Kavlock RJ, et al. (2011). Predictive model of rat reproductive toxicity from ToxCast high throughput screening. *Biology of Reproduction* 85: 327-339. http://www.ncbi.nlm.nih.gov/pubmed/21565999.
Martin MT, Mendez E, Corum DG, Judson RS, Kavlock RJ, Rotroff DM, et al. (2009b). Profiling the reproductive toxicity of chemicals from multigeneration studies in the toxicity reference database. *Toxicological Sciences* 110: 181-190. http://www.ncbi.nlm.nih.gov/pubmed/19363143.
Maull EA, Ahsan H, Edwards J, Longnecker MP, Navas-Acien A, Pi J, et al. (2012). Evaluation of the association between arsenic and diabetes: A National Toxicology Program workshop review. *Environmental Health Perspectives* 120: 1658-1670. http://www.ncbi.nlm.nih.gov/pubmed/22889723.
Mayr LM, Bojanic D. (2009). Novel trends in high-throughput screening. *Current Opinion in Pharmacology* 9: 580-588. http://www.ncbi.nlm.nih.gov/pubmed/19775937.
McCullough SD, Duncan KE, Swanton SM, Dailey LA, Diaz-Sanchez D, Devlin RB. (in press). Ozone induces a pro-inflammatory response in primary human bronchial epithelial cells through mitogen-activated protein kinase activation without nuclear factor-κB activation. *Am J Respir Cell Mol Biol*. [Epub ahead of print]. http://www.ncbi.nlm.nih.gov/pubmed/24693990.
McDonnell WF. (1996). Individual variability in human lung function responses to ozone exposure. *Environ Toxicol Pharmacol* 2: 171-175. http://www.ncbi.nlm.nih.gov/pubmed/21781722.
McHale CM, Zhang L, Lan Q, Li G, Hubbard AE, Forrest MS, et al. (2009). Changes in the peripheral blood transcriptome associated with occupational benzene exposure identified by cross-comparison on two microarray platforms. *Genomics* 93: 343-349. http://www.ncbi.nlm.nih.gov/pubmed/19162166.
McHale CM, Zhang L, Lan Q, Vermeulen R, Li G, Hubbard AE, et al. (2010). Global gene expression profiling of a population exposed to a range of benzene levels. *Environmental Health Perspectives* 119: 628-634. http://dx.doi.org/10.1289/ehp.1002546.
McHale CM, Zhang L, Smith MT. (2012). Current understanding of the mechanism of benzene-induced leukemia in humans: Implications for risk assessment. *Carcinogenesis* 33: 240-252. http://www.ncbi.nlm.nih.gov/pubmed/22166497.
McLanahan ED, El-Masri HA, Sweeney LM, Kopylev LY, Clewell HJ, Wambaugh JF, et al. (2012). Physiologically based pharmacokinetic model use in risk assessment--Why being published is not enough. *Toxicological Sciences* 126: 5-15. http://www.ncbi.nlm.nih.gov/pubmed/22045031.
Mechanic LE, Chen HS, Amos CI, Chatterjee N, Cox NJ, Divi RL, et al. (2012). Next generation analytic tools for large scale genetic epidemiology studies of complex diseases. *Genetic Epidemiology* 36: 22-35. http://www.ncbi.nlm.nih.gov/pubmed/22147673.
Medzhitov R. (2008). Origin and physiological roles of inflammation. *Nature* 454: 428-435. http://www.ncbi.nlm.nih.gov/pubmed/18650913.
Meek ME, Boobis A, Cote I, Dellarco V, Fotakis G, Munn S, et al. (2014). New developments in the evolution and application of the WHO/IPCS framework on mode of action/species concordance analysis. *J Appl Toxicol* 34: 1-18. http://onlinelibrary.wiley.com/doi/10.1002/jat.2949/abstract.
Meek ME, Bucher JR, Cohen SM, Dellarco V, Hill RN, Lehman-McKeeman LD, et al. (2003). A framework for human relevance analysis of information on carcinogenic modes of action. *Critical Reviews in Toxicology* 33: 591-653.

http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=14727733.
Meissner A. (2012). What can epigenomics do for you? *Genome Biology* 13: 420. http://www.ncbi.nlm.nih.gov/pubmed/23095436.
Mendrick DL. (2011). Transcriptional profiling to identify biomarkers of disease and drug response. *Pharmacogenomics* 12: 235-249. http://www.ncbi.nlm.nih.gov/pubmed/21332316.
Miao X, Sun W, Fu Y, Miao L, Cai L. (2013). Zinc homeostasis in the metabolic syndrome and diabetes. *Frontiers of Medicine* 7: 31-52. http://www.ncbi.nlm.nih.gov/pubmed/23385610.
Miller MD, Crofton KM, Rice DC, Zoeller RT. (2009). Thyroid-disrupting chemicals: Interpreting upstream biomarkers of adverse outcomes. *Environmental Health Perspectives* 117: 1033-1041. http://www.ncbi.nlm.nih.gov/pubmed/19654909.
Mitchell J, Arnot JA, Jolliet O, Georgopoulos PG, Isukapalli S, Dasgupta S, et al. (2013). Comparison of modeling approaches to prioritize chemicals based on estimates of exposure and exposure potential. *Sci Total Environ* 458-460: 555-567. http://www.ncbi.nlm.nih.gov/pubmed/23707726.
Mitra S, Das S, Chakrabarti J. (2013). Systems biology of cancer biomarker detection. *Cancer Biomark* 13: 201-213. http://www.ncbi.nlm.nih.gov/pubmed/24240581.
Molinelli EJ, Korkut A, Wang W, Miller ML, Gauthier NP, Jing X, et al. (2013). Perturbation biology: Inferring signaling networks in cellular systems. *PLoS Comput Biol* 9: e1003290. http://www.ncbi.nlm.nih.gov/pubmed/24367245.
Morales-Ruan MC, Villalpando S, Garcia-Guerra A, Shamah-Levy T, Robledo-Perez R, Avila-Arcos MA, et al. (2012). Iron, zinc, copper and magnesium nutritional status in Mexican children aged 1 to 11 years. *Salud Publica de Mexico* 54: 125-134. http://www.ncbi.nlm.nih.gov/pubmed/22535171.
Moraru, II, Schaff JC, Slepchenko BM, Blinov ML, Morgan F, Lakshminarayana A, et al. (2008). Virtual cell modelling and simulation software environment. *IET Systems Biology* 2: 352-362. http://www.ncbi.nlm.nih.gov/pubmed/19045830.
Mortensen HM, Euling SY. (2013). Integrating mechanistic and polymorphism data to characterize human genetic susceptibility for environmental chemical risk assessment in the 21st century. *Toxicology and Applied Pharmacology* 271: 395-404. http://dx.doi.org/10.1016/j.taap.2011.01.015.
Mumtaz M, Fisher J, Blount B, Ruiz P. (2012). Application of physiologically based pharmacokinetic models in chemical risk assessment. *Journal of Toxicology* 2012: 11. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3317240/.
Murk AJ, Rijntjes E, Blaauboer BJ, Clewell R, Crofton KM, Dingemans MM, et al. (2013). Mechanism-based testing strategy using *in vitro* approaches for identification of thyroid hormone disrupting chemicals. *Toxicology In Vitro*. http://www.ncbi.nlm.nih.gov/pubmed/23453986.
NAFTA (North American Free Trade Agreement). (2012). *(Quantitative) Structure Activity Relationship [(Q)SAR] guidance document*. Technical Working Group on Pesticides (TWG). Retrieved from http://www.epa.gov/oppfead1/international/naftatwg/guidance/qsar-guidance.pdf.
NCBI (National Center for Biotechnology Information). (2012). *Reference SNP(refSNP) Cluster Report: rs13266634*. dbSNP: Short Genetic Variations. Retrieved from http://www.ncbi.nlm.nih.gov/projects/SNP/snp_ref.cgi?rs=13266634 (accessed March 20, 2013).
NCBI (National Center for Biotechnology Information). (2014a). *BioSystems Database*. Retrieved from http://www.ncbi.nlm.nih.gov/Structure/biosystems/docs/biosystems_about.html (accessed August 29, 2014).

NCBI (National Center for Biotechnology Information). (2014b). *Epigenomics (database)*. Retrieved from http://www.ncbi.nlm.nih.gov/epigenomics (accessed August 29, 2014).
NCBI (National Center for Biotechnology Information). (2014c). *Gene Expression Omnibus*. Retrieved from http://www.ncbi.nlm.nih.gov/geo/ (accessed August 29, 2014).
NCBI (National Center for Biotechnology Information). (2014d). *NCBI All Resources*. Retrieved from http://www.ncbi.nlm.nih.gov/guide/all/ (accessed August 29, 2014).
NCBI (National Center for Biotechnology Information). (2014e). *PubChem BioAssay Database*. Retrieved from http://www.ncbi.nlm.nih.gov/pcassay (accessed August 29, 2014).
NHGRI (National Human Genome Research Institute). (2014a). *A Catalog of Published Genome-Wide Association Studies*. Retrieved from http://www.genome.gov/gwastudies/ (accessed August 29, 2014).
NHGRI (National Human Genome Research Institute). (2014b). *Home Page*. Retrieved from http://www.genome.gov/ (accessed August 29, 2014).
Nichols JW, Breen M, Denver RJ, Distefano JJ, 3rd, Edwards JS, Hoke RA, et al. (2011). Predicting chemical impacts on vertebrate endocrine systems. *Environmental Toxicology and Chemistry* 30: 39-51. http://www.ncbi.nlm.nih.gov/pubmed/20963851.
Nie AY, Mcmillian M, Parker JB, Leone A, Bryant S, Yieh L, et al. (2006). Predictive toxicogenomics approaches reveal underlying molecular mechanisms of nongenotoxic carcinogenicity. *Molecular Carcinogenesis* 45: 914-933. http://www.ncbi.nlm.nih.gov/pubmed/16921489.
NIEHS (National Institute of Environmental Health Sciences). (2014a). *Comparative Toxicogenomic Database (CTD)™*. Retrieved from http://ctdbase.org/ (accessed August 29, 2014).
NIEHS (National Institute of Environmental Health Sciences). (2014b). *Host Susceptibility Program*. Retrieved from http://grants.nih.gov/grants/guide/notice-files/NOT-ES-07-009.html (accessed August 29, 2014).
NIEHS (National Institute of Environmental Health Sciences). (2014c). *Mapping the Human Toxome by Systems Toxicology. Grant Number R01ES020750*. Retrieved from http://tools.niehs.nih.gov/portfolio/index.cfm/portfolio/grantDetail/grant_number/R01ES020750 (accessed August 29, 2014).
NIEHS (National Institute of Environmental Health Sciences). (2014d). *NIEHS SNPs Environmental Genome Project*. Retrieved from http://egp.gs.washington.edu/ (accessed August 29, 2014).
NIEHS (National Institute of Environmental Health Sciences). (2014e). *NIEHS Strategic Plan*. Retrieved from http://niehs.nih.gov/about/strategicplan/index.cfm (accessed August 29, 2014).
NIH (National Institutes of Health). (2014). *NIH Chemical Genomics Center*. Retrieved from http://www.ncats.nih.gov/research/reengineering/ncgc/ncgc.html (accessed August 29, 2014).
NIOSH (National Institute of Occupational Safety and Health). (1992). *Recommendations for Occupational Safety and Health: Compendium of Policy Documents and Statements*. Retrieved from http://www.cdc.gov/niosh/pdfs/92-100.pdf (accessed August 29, 2014).
North M, Tandon VJ, Thomas R, Loguinov A, Gerlovina I, Hubbard AE, et al. (2011). Genome-wide functional profiling reveals genes required for tolerance to benzene metabolites in yeast. *Public Library of Science One* 6: e24205. http://www.ncbi.nlm.nih.gov/pubmed/21912624.
NRC (National Research Council). (2005). *Application of Toxicogenomics to Cross-Species Extrapolation:A Report of a Workshop*. (9780309100847). The National Academies Press. Retrieved from http://www.nap.edu/openbook.php?record_id=11488.
NRC (National Research Council). (2006). *Human Biomonitoring for Environmental Chemicals*. The National Academies Press. Retrieved from http://www.nap.edu/openbook.php?record_id=11700.
NRC (National Research Council). (2007a). *Scientific Review of the Proposed Risk Assessment Bulletin from the Office of Management and Budget*. (9780309104777). Washington, DC: The

National Academies Press. Retrieved from http://www.nap.edu/openbook.php?record_id=11811.
NRC (National Research Council). (2007b). *Toxicity Testing in the 21st Century: A Vision and a Strategy*. Washington DC. Retrieved from http://dels.nas.edu/resources/static-assets/materials-based-on-reports/reports-in-brief/Toxicity_Testing_final.pdf (accessed August 29, 2014).
NRC (National Research Council). (2009). *Science and Decisions: Advancing Risk Assessment*. Washington, DC. Retrieved from http://www.nap.edu/catalog/12209.html (accessed August 29, 2014).
NRC (National Research Council). (2011). *Predicting Later-Life Outcomes of Early-Life Exposures*. Retrieved from http://nas-sites.org/emergingscience/files/2011/05/inutero_final_April2011.pdf (accessed August 29, 2014).
NRC (National Research Council). (2013). *Environmental Decisions in the Face of Uncertainty*. (9780309130349). The National Academies Press. Retrieved from http://www.nap.edu/openbook.php?record_id=12568.
NRC (National Research Council). (2014). *Review of EPA's Integrated Risk Information System (IRIS) Process*. Washington, D.C.: The National Academies Press. Retrieved from http://www.nap.edu/catalog.php?record_id=18764.
O'Shea SH, Schwarz J, Kosyk O, Ross PK, Ha MJ, Wright FA, et al. (2011). *In vitro* screening for population variability in chemical toxicity. *Toxicological Sciences* 119: 398-407. http://www.ncbi.nlm.nih.gov/pubmed/20952501.
OECD (Organization for Economic Cooperation and Development). (2004a). *Detailed review raper on amphibian metamorphosis assays for the detection of thyroid active substances*. (ENV/JM/MONO(2004)17). Retrieved from http://www.epa.gov/endo/pubs/attachment_e_oecd_amphibian_drp.pdf (accessed August 29, 2014).
OECD (Organization for Economic Cooperation and Development). (2004b). *OECD Principles for the Validation, for Regulatory Purposes, of Quantitative Structure Activity Relationship Models*. Paris, France. Retrieved from http://www.oecd.org/chemicalsafety/assessmentofchemicals/37849783.pdf (accessed FAugust 29, 2014).
OECD (Organization for Economic Cooperation and Development). (2010). *Other activities on molecular screening and toxicogenomics*. Retrieved from http://www.oecd.org/env/ehs/testing/toxicogenomics.htm (accessed August 29, 2014).
OECD (Organization for Economic Cooperation and Development). (2013). *Guidance document on developing and assessing adverse outcome pathways*. Retrieved from http://search.oecd.org/officialdocuments/displaydocumentpdf/?cote=env/jm/mono%282013%296&doclanguage=en (accessed August 29, 2014).
OECD (Organization for Economic Cooperation and Development). (2014a). *Adverse outcome pathways, molecular screening and toxicogenomics*. Retrieved from http://www.oecd.org/env/ehs/testing/adverse-outcome-pathways-molecular-screening-and-toxicogenomics.htm (accessed August 29, 2014).
OECD (Organization for Economic Cooperation and Development). (2014b). *Grouping of chemicals: Chemical categories and read-across*. Retrieved from http://www.oecd.org/chemicalsafety/risk-assessment/groupingofchemicalschemicalcategoriesandread-across.htm (accessed August 29, 2014).
OECD (Organization for Economic Cooperation and Development). (2014c). *The OECD QSAR Toolbox. Version 3.0*. Retrieved from

http://www.oecd.org/env/chemicalsafetyandbiosafety/assessmentofchemicals/theoecdqsartoolbox.htm (accessed August 29, 2014).
OECD (Organization for Economic Cooperation and Development). (2014d). *Other activities on molecular screening and toxicogenomics*. Retrieved from http://www.oecd.org/env/ehs/testing/toxicogenomics.htm (accessed August 29, 2014).
OECD (Organization for Economic Cooperation and Development). (2014e). *QSAR toolbox for grouping chemicals into categories*. Retrieved from http://www.qsartoolbox.org/home (accessed August 29, 2014).
OMB/OSTP (Office of Management and Budget/Office of Science and Technology Policy). (2007). *OMB and OSTP Statements on Updated 1995 Principles for Risk Analysis*. Retrieved from http://georgewbush-whitehouse.archives.gov/omb/pubpress/2007/091907_risk_analysis.html.
OMIM (Online Mendelian Inheritance in Man). (2014). *Solute Carrier Family 30 (Zinc Transporter), Member 8, SLC30A8* Retrieved from http://omim.org/entry/611145#0001 (accessed August 29, 2014).
OSHA (Occupational Safety & Health Administration). (2014). *Chemical sampling information: Benzene*. Retrieved from https://www.osha.gov/dts/chemicalsampling/data/CH_220100.html (accessed
Padilla S, Corum D, Padnos B, Hunter DL, Beam A, Houck KA, et al. (2012). Zebrafish developmental screening of the ToxCast Phase I chemical library. *Reproductive Toxicology* 33: 174-187. http://www.ncbi.nlm.nih.gov/pubmed/22182468.
Pare G, Chasman DI, Parker AN, Nathan DM, Miletich JP, Zee RY, et al. (2008). Novel association of HK1 with glycated hemoglobin in a non-diabetic population: A genome-wide evaluation of 14,618 participants in the Women's Genome Health Study. *Public Library of Science Genetics* 4: e1000312. http://www.ncbi.nlm.nih.gov/pubmed/19096518.
Parham F, Austin C, Southall N, Huang R, Tice R, Portier C. (2009). Dose-response modeling of high-throughput screening data. *Journal of Biomolecular Screening* 14: 1216-1227. http://www.ncbi.nlm.nih.gov/pubmed/19828774.
Paris M, Laudet V. (2008). The history of a developmental stage: Metamorphosis in chordates. *Genesis* 46: 657-672. http://www.ncbi.nlm.nih.gov/pubmed/18932261.
Parng C, Seng WL, Semino C, McGrath P. (2002). Zebrafish: A preclinical model for drug screening. *Assay & Drug Development Technologies* 1: 41-48. http://www.ncbi.nlm.nih.gov/pubmed/15090155.
Patel CJ, Chen R, Butte AJ. (2012a). Data-driven integration of epidemiological and toxicological data to select candidate interacting genes and environmental factors in association with disease. *Bioinformatics* 28: i121-126. http://www.ncbi.nlm.nih.gov/pubmed/22689751.
Patel CJ, Chen R, Kodama K, Loannidis JP, Butte AJ. (2013). Systematic identification of interaction effects between genome-and environment-wide associations in type 2 diabetes mellitus. *Human Genetics*. http://www.ncbi.nlm.nih.gov/pubmed/23334806.
Patel CJ, Cullen MR. (2012). Genetic variability in molecular responses to chemical exposure. *Experientia* 101: 437-457. http://www.ncbi.nlm.nih.gov/pubmed/22945578.
Patel CJ, Cullen MR, Loannidis JP, Butte AJ. (2012b). Systematic evaluation of environmental factors: Persistent pollutants and nutrients correlated with serum lipid levels. *International Journal of Epidemiology* 41: 828-843. http://dx.doi.org/10.1093/ije/dys003.
Patlewicz G, Ball N, Booth ED, Hulzebos E, Zvinavashe E, Hennes C. (2013a). Use of category approaches, read-across and (Q)SAR: general considerations. *Regul Toxicol Pharmacol* 67: 1-12. http://www.ncbi.nlm.nih.gov/pubmed/23764304.
Patlewicz G, Roberts DW, Aptula A, Blackburn K, Hubesch B. (2013b). Workshop: Use of "read-across" for chemical safety assessment under REACH. *Regul Toxicol Pharmacol* 65: 226-228. http://www.sciencedirect.com/science/article/pii/S0273230012002450.

Peden DB, Boehlecke B, Horstman D, Devlin R. (1997). Prolonged acute exposure to 0.16 ppm ozone induces eosinophilic airway inflammation in asthmatic subjects with allergies. *Journal of Allergy and Clinical Immunology* 100: 802-808. http://www.ncbi.nlm.nih.gov/pubmed/9438490.

Pedersen-Bjergaard J, Andersen MK, Andersen MT, Christiansen DH. (2008). Genetics of therapy-related myelodysplasia and acute myeloid leukemia. *Leukemia* 22: 240-248. http://www.ncbi.nlm.nih.gov/pubmed/18200041.

Perkins EJ, Ankley GT, Crofton KM, Garcia-Reyero N, Lalone CA, Johnson MS, et al. (2013). Current perspectives on the use of alternative species in human health and ecological risk assessments. *Environmental Health Perspectives*. http://www.ncbi.nlm.nih.gov/pubmed/23771518.

Peyret T, Krishnan K. (2011). QSARs for PBPK modelling of environmental contaminants. *SAR QSAR Environ Res* 22: 129-169. http://www.ncbi.nlm.nih.gov/pubmed/21391145.

Physiome Project. (2014). *Physiome Project Website*. Retrieved from http://physiomeproject.org/ (accessed August 29, 2014).

Pilsner JR, Liu X, Ahsan H, Ilievski V, Slavkovich V, Levy D, et al. (2009). Folate deficiency, hyperhomocysteinemia, low urinary creatinine, and hypomethylation of leukocyte DNA are risk factors for arsenic-induced skin lesions. *Environ Health Perspect* 117: 254-260. http://www.ncbi.nlm.nih.gov/pubmed/19270796.

Poulin P, Haddad S. (2013). Toward a new paradigm for the efficient *in vitro-in vivo* extrapolation of metabolic clearance in humans from hepatocyte data. *J Pharm Sci* 102: 3239-3251. http://www.ncbi.nlm.nih.gov/pubmed/23494893.

Raaijmakers MH. (2011). Niche contributions to oncogenesis: Emerging concepts and implications for the hematopoietic system. *Haematologica* 96: 1041-1048. http://www.ncbi.nlm.nih.gov/pubmed/21459792.

Rabinowitz JR, Goldsmith MR, Little SB, Pasquinelli MA. (2008). Computational molecular modeling for evaluating the toxicity of environmental chemicals: Prioritizing bioassay requirements. *Environmental Health Perspectives* 116: 573-577. http://www.ncbi.nlm.nih.gov/pubmed/18470285.

Rakyan VK, Down TA, Balding DJ, Beck S. (2011). Epigenome-wide association studies for common human diseases. *Nature Reviews Genetics* 12: 529-541. http://www.ncbi.nlm.nih.gov/pubmed/21747404.

Raldua D, Thienpont B, Babin PJ. (2012). Zebrafish eleutheroembryos as an alternative system for screening chemicals disrupting the mammalian thyroid gland morphogenesis and function. *Reproductive Toxicology* 33: 188-197. http://www.ncbi.nlm.nih.gov/pubmed/21978863.

Ramasamy A, Mondry A, Holmes CC, Altman DG. (2008). Key issues in conducting a meta-analysis of gene expression microarray datasets. *Public Library of Science Medicine* 5: e184. http://www.ncbi.nlm.nih.gov/pubmed/18767902.

REACH. (2014). *EU's - Registration, Evaluation, Authorisation and Restriction of Chemicals*. Retrieved from http://echa.europa.eu/support/grouping-of-substances-and-read-across (accessed August 29, 2014).

Reaume CJ, Sokolowski MB. (2011). Conservation of gene function in behaviour. *Philosophical Transactions of the Royal Society London B: Biological Science* 366: 2100-2110. http://www.ncbi.nlm.nih.gov/pubmed/21690128.

Reif DM, Martin MT, Tan SW, Houck KA, Judson RS, Richard AM, et al. (2010). Endocrine profiling and prioritization of environmental chemicals using ToxCast data. *Environmental Health Perspectives* 118: 1714-1720. http://www.ncbi.nlm.nih.gov/pubmed/20826373.

Reuschenbach P, Silvani M, Dammann M, Warnecke D, Knacker T. (2008). ECOSAR model performance with a large test set of industrial chemicals. *Chemosphere* 71: 1986-1995. http://www.ncbi.nlm.nih.gov/pubmed/18262586.

Rhomberg LR, Goodman JE, Bailey LA, Prueitt RL, Beck NB, Bevan C, et al. (2013). A survey of frameworks for best practices in weight-of-evidence analyses. *Critical Reviews in Toxicology* 43: 753-784. http://informahealthcare.com/doi/abs/10.3109/10408444.2013.832727.

Rhomberg LR, Goodman JE, Haber LT, Dourson M, Andersen ME, Klaunig JE, et al. (2011). Linear low-dose extrapolation for noncancer heath effects is the exception, not the rule. *Crit Rev Toxicol* 41: 1-19. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3038594/.

Rosenbaum RK, Bachmann TM, Swirsky Gold L, Huijbregts MAJ, Jolliet O, Juraske R. (2008). USEtox -- the UNEP-SETAC toxicity model: Recommended characterization factors for human toxicity and freshwater ecotoxicicty in life cycle impact assessment. *International Journal of Life Cycle Assessment* 13: 532-546. http://link.springer.com/article/10.1007%2Fs11367-008-0038-4#.

Rotroff DM, Dix DJ, Houck KA, Kavlock RJ, Knudsen TB, Martin MT, et al. (2013). Real-time growth kinetics measuring hormone mimicry for ToxCast chemicals in T-47D human ductal carcinoma cells. *Chem Res Toxicol* 26: 1097-1107. http://www.ncbi.nlm.nih.gov/pubmed/23682706.

Rotroff DM, Dix DJ, Houck KA, Knudsen TB, Martin MT, McLaurin KW, et al. (2012). Using *in vitro* high throughput screening assays to identify potential endocrine-disrupting chemicals. *Environmental Health Perspectives* Epub doi: 10.1289/ehp.1205065. http://www.ncbi.nlm.nih.gov/pubmed/23052129.

Rotroff DM, Wetmore BA, Dix DJ, Ferguson SS, Clewell HJ, Houck KA, et al. (2010). Incorporating human dosimetry and exposure into high-throughput *in vitro* toxicity screening. *Toxicological Sciences* 117: 348-358. http://www.ncbi.nlm.nih.gov/pubmed/20639261.

Rudel RA, Dodson RW, Newton E, Zota AR, Brody JG. (2008). Correlations between urinary phthalate metabolites and phthalates, estrogenic compounds 4-butyl phenol and o-phenyl phenol, and some pesticides in home indoor air and house dust. *Epidemiology* 19: S332. http://www.researchgate.net/publication/232131801_Correlations_Between_Urinary_Phthalate_Metabolites_and_Phthalates_Estrogenic_Compounds_4-Butyl_phenol_and_o-Phenyl_phenol_and_Some_Pesticides_in_Home_Indoor_Air_and_House_Dust.

Rung J, Cauchi S, Albrechtsen A, Shen L, Rocheleau G, Cavalcanti-Proenca C, et al. (2009). Genetic variant near IRS1 is associated with type 2 diabetes, insulin resistance and hyperinsulinemia. *Nature Genetics* 41: 1110-1115. http://www.ncbi.nlm.nih.gov/pubmed/19734900.

Rusyn I, Gatti DM, Wiltshire T, Kleeberger SR, Threadgill DW. (2010). Toxicogenetics: Population-based testing of drug and chemical safety in mouse models. *Pharmacogenomics* 11: 1127-1136. http://www.ncbi.nlm.nih.gov/pubmed/20704464.

Rusyn I, Sedykh A, Low Y, Guyton KZ, Tropsha A. (2012). Predictive modeling of chemical hazard by integrating numerical descriptors of chemical structures and short-term toxicity assay data. *Toxicological Sciences* 127: 1-9. http://dx.doi.org/10.1093/toxsci/kfs095.

SAB (Science Advisory Board). (2013). *Draft SAB advice on advancing the application of computational toxicology research for human health risk assessment*. Washington, DC. Retrieved from http://yosemite.epa.gov/sab/sabproduct.nsf/46963ceebabd621905256cae0053d5c6/F3315D0EE2EDC11285257B0A005F5B7E/$File/CompTox-edited+1-29-13.pdf (accessed August 29, 2014).

Sagredo C, Ovrebo S, Haugen A, Fujii-Kuriyama Y, Baera R, Botnen IV, et al. (2006). Quantitative analysis of benzo[a]pyrene biotransformation and adduct formation in Ahr knockout mice. *Toxicology Letters* 167: 173-182. http://www.ncbi.nlm.nih.gov/pubmed/17049425.

Samuels MH, Luther M, Henry P, Ridgway EC. (1994). Effects of hydrocortisone on pulsatile pituitary glycoprotein secretion. *Journal of Clinical Endocrinology & Metabolism* 78: 211-215. http://www.ncbi.nlm.nih.gov/pubmed/8288706.

Sand S, Portier CJ, Krewski D. (2011). A signal-to-noise crossover dose as the point of departure for health risk assessment. *Environmental Health Perspectives* 119: 1766-1774. http://www.ncbi.nlm.nih.gov/pubmed/21813365.

Sarapura VD, Samuels MH, Ridgway EC. (2002). Thyroid-Stimulating Hormone. In S Melmed (Ed.), *The Pitutary* (Second ed., pp. 187-229). Malden, MA: Blackwell Science.

Schnatter AR, Glass DC, Tang G, Irons RD, Rushton L. (2012). Myelodysplastic syndrome and benzene exposure among petroleum workers: An international pooled analysis. *Journal of the National Cancer Institute* 104: 1724-1737. http://www.ncbi.nlm.nih.gov/pubmed/23111193.

Schreiber R, Gundel U, Franz S, Kuster A, Rechenberg B, Altenburger R. (2011). Using the fish plasma model for comparative hazard identification for pharmaceuticals in the environment by extrapolation from human therapeutic data. *Regulatory Toxicology and Pharmacology* 61: 261-275. http://www.ncbi.nlm.nih.gov/pubmed/21889559.

Schreiber S. (2003). The small-molecule approach to biology: Chemical genetics and diversity-oriented organic synthesis make possible the systematic exploration of biology. *Chemical & Engineering News* 81: 51-61. http://pubs.acs.org/cen/coverstory/8109/8109genomics.html.

Schug TT, Janesick A, Blumberg B, Heindel JJ. (2011). Endocrine disrupting chemicals and disease susceptibility. *Journal of Steroid Biochemistry and Molecular Biology* 127: 204-215. http://www.ncbi.nlm.nih.gov/pubmed/21899826.

Scott LJ, Mohlke KL, Bonnycastle LL, Willer CJ, Li Y, Duren WL, et al. (2007). A genome-wide association study of type 2 diabetes in Finns detects multiple susceptibility variants. *Science* 316: 1341-1345. http://www.ncbi.nlm.nih.gov/pubmed/17463248.

Sedykh A, Zhu H, Tang H, Zhang L, Richard A, Rusyn I, et al. (2011). Use of *in vitro* HTS-derived concentration-response data as biological descriptors improves the accuracy of QSAR models of *in vivo* toxicity. *Environmental Health Perspectives* 119: 364-370. http://www.ncbi.nlm.nih.gov/pubmed/20980217.

Seed J, Carney EW, Corley RA, Crofton KM, DeSesso JM, Foster PM, et al. (2005). Overview: Using mode of action and life stage information to evaluate the human relevance of animal toxicity data. *Critical Reviews in Toxicology* 35: 664-672. http://www.ncbi.nlm.nih.gov/pubmed/16417033.

Selgrade MK, Cooper KD, Devlin RB, van Loveren H, Biagini RE, Luster MI. (1995). Immunotoxicity--bridging the gap between animal research and human health effects. *Fundamental and Applied Toxicology* 24: 13-21. http://www.ncbi.nlm.nih.gov/pubmed/7713335.

Serafimova R, Todorov M, Nedelcheva D, Pavlov T, Akahori Y, Nakai M, et al. (2007). QSAR and mechanistic interpretation of estrogen receptor binding. *SAR and QSAR in Environmental Research* 18: 389-421. http://www.ncbi.nlm.nih.gov/pubmed/17514577.

Shaffer CL, Scialis RJ, Rong H, Obach RS. (2012). Using Simcyp to project human oral pharmacokinetic variability in early drug research to mitigate mechanism-based adverse events. *Biopharmaceutics & Drug Disposition* 33: 72-84. http://www.ncbi.nlm.nih.gov/pubmed/22213407.

Shah I, Houck K, Judson R, Kavlock RJ, Martin M, Reif DM, et al. (2011). Using nuclear receptor activity to stratify hepatocarcinogens. *PloS One* 6. http://www.plosone.org/article/info%3Adoi%2F10.1371%2Fjournal.pone.0014584.

Shah I, Wambaugh J. (2010). Virtual tissues in toxicology. *Journal of Toxicology and Environmental Health Part B: Critical Reviews* 13: 314-328. http://www.ncbi.nlm.nih.gov/pubmed/20574905.

Sheldon LS, Cohen Hubal EA. (2009). Exposure as part of a systems approach for assessing risk. *Environmental Health Perspectives* 117: 119-1194. http://www.ncbi.nlm.nih.gov/pubmed/19672394.

Shen M, Zhang L, Lee KM, Vermeulen R, Hosgood HD, Li G, et al. (2011). Polymorphisms in genes involved in innate immunity and susceptibility to benzene-induced hematotoxicity. *Experimental & Molecular Medicine* 43: 374-378. http://www.ncbi.nlm.nih.gov/pubmed/21540635.

Shi L, Jones WD, Jensen RV, Harris SC, Perkins RG, Goodsaid FM, et al. (2008). The balance of reproducibility, sensitivity, and specificity of lists of differentially expressed genes in microarray studies. *BMC Bioinformatics* 9 Suppl 9: S10. http://www.ncbi.nlm.nih.gov/pubmed/18793455.

Sille FC, Thomas R, Smith MT, Conde L, Skibola CF. (2012). Post-GWAS functional characterization of susceptibility variants for chronic lymphocytic leukemia. *Public Library of Science One* 7: e29632. http://www.ncbi.nlm.nih.gov/pubmed/22235315.

Sipes NS, Martin MT, Kothiya P, Reif DM, Judson RS, Richard AM, et al. (2013). Profiling 976 ToxCast chemicals across 331 enzymatic and receptor signaling assays. *Chemical Research in Toxicology* 26: 878-895. http://www.ncbi.nlm.nih.gov/pubmed/23611293.

Sipes NS, Martin MT, Reif DM, Kleinstreuer NC, Judson RS, Singh AV, et al. (2011a). Predictive models of prenatal developmental toxicity from ToxCast high-throughput screening data. *Toxicological Sciences* 124: 109-127. http://toxsci.oxfordjournals.org/content/124/1/109.abstract.

Sipes NS, Padilla S, Knudsen TB. (2011b). Zebrafish: As an integrative model for twenty-first century toxicity testing. *Birth Defects Research: Part C, Embryo Today* 93: 256-267. http://www.ncbi.nlm.nih.gov/pubmed/21932434.

Sistonen J, Sajantila A, Lao O, Corander J, Barbujani G, Fuselli S. (2007). CYP2D6 worldwide genetic variation shows high frequency of altered activity variants and no continental structure. *Pharmacogenetics and Genomics* 17: 93-101. http://www.ncbi.nlm.nih.gov/pubmed/17301689.

Skolness SY, Blanksma CA, Cavallin JE, Churchill JJ, Durhan EJ, Jensen KM, et al. (2013). Propiconazole inhibits steroidogenesis and reproduction in the fathead minnow (Pimephales promelas). *Toxicological Sciences*. http://www.ncbi.nlm.nih.gov/pubmed/23339182.

Sladek R, Rocheleau G, Rung J, Dina C, Shen L, Serre D, et al. (2007). A genome-wide association study identifies novel risk loci for type 2 diabetes. *Nature* 445: 881-885. http://www.ncbi.nlm.nih.gov/pubmed/17293876.

Slob W, Bakker MI, Biesebeek JD, Bokkers BG. (in press). Exploring the uncertainties in cancer risk assessment using the integrated probabilistic risk assessment (IPRA) approach. *Risk Anal*. [Epub ahead of print]. http://www.ncbi.nlm.nih.gov/pubmed/24766324.

Smirnova L, Hogberg HT, Leist M, Hartung T. (2014). Developmental neurotoxicity - challenges in the 21st century and *in vitro* opportunities. *ALTEX* 31: 129-156. http://www.ncbi.nlm.nih.gov/pubmed/24687333.

Smith MT, Zhang L, McHale CM, Skibola CF, Rappaport SM. (2011). Benzene, the exposome and future investigations of leukemia etiology. *Chemico-Biological Interactions* 192: 155-159. http://www.ncbi.nlm.nih.gov/pubmed/21333640.

Smith MV, Boyd WA, Kissling GE, Rice JR, Snyder DW, Portier CJ, et al. (2009). A discrete time model for the analysis of medium-throughput C. elegans growth data. *Public Library of Science One* 4: e7018. http://www.ncbi.nlm.nih.gov/pubmed/19753303.

Sonich-Mullin C, Fielder R, Wiltse J, Baetcke K, Dempsey J, Fenner-Crisp P, et al. (2001). IPCS conceptual framework for evaluating a mode of action for chemical carcinogenesis. *Regulatory Toxicology and Pharmacology* 34: 146-152. http://www.ncbi.nlm.nih.gov/pubmed/11603957.

Steinthorsdottir V, Thorleifsson G, Reynisdottir I, Benediktsson R, Jonsdottir T, Walters GB, et al. (2007). A variant in CDKAL1 influences insulin response and risk of type 2 diabetes. *Nature Genetics* 39: 770-775. http://www.ncbi.nlm.nih.gov/pubmed/17460697.

Sturla SJ, Boobis AR, Fitzgerald RE, Hoeng J, Kavlock RJ, Schirmer K, et al. (2014). Systems toxicology: From basic research to risk assessment. *Chem Res Toxicol* 27: 314-329. http://www.ncbi.nlm.nih.gov/pubmed/24446777.

Swat MH, Thomas GL, Belmonte JM, Shirinifard A, Hmeljak D, Glazier JA. (2012). Chapter 13 - Multi-Scale Modeling of Tissues Using CompuCell3D. In *Methods in Cell Biology*: Academic Press. http://www.sciencedirect.com/science/article/pii/B9780123884039000138.

Takeuchi F, Serizawa M, Yamamoto K, Fujisawa T, Nakashima E, Ohnaka K, et al. (2009). Confirmation of multiple risk Loci and genetic impacts by a genome-wide association study of type 2 diabetes in the Japanese population. *Diabetes* 58: 1690-1699. http://www.ncbi.nlm.nih.gov/pubmed/19401414.

Taylor KW, Novak RF, Anderson HA, Birnbaum LS, Blystone C, DeVito M, et al. (2013). Evaluation of the Association between Persistent Organic Pollutants (POPs) and Diabetes in Epidemiological Studies: A National Toxicology Program Workshop Review. *Environmental Health Perspectives* 121: 774-783. http://www.ncbi.nlm.nih.gov/pubmed/23651634.

Teschendorff AE, Widschwendter M. (2012). Differential variability improves the identification of cancer risk markers in DNA methylation studies profiling precursor cancer lesions. *Bioinformatics* 28: 1487-1494. http://www.ncbi.nlm.nih.gov/pubmed/22492641.

Thayer KA, Heindel JJ, Bucher JR, Gallo MA. (2012). Role of environmental chemicals in diabetes and obesity: A National Toxicology Program Workshop report. *Environmental Health Perspectives* 120: 779-789. http://dx.doi.org/10.1289/ehp.1104597.

The 1000 Genomes Project Consortium. (2010). A map of human genome variation from population-scale sequencing. *Nature* 467: 1061-1073. http://www.1000genomes.org/sites/1000genomes.org/files/docs/nature09534.pdf.

The Hamner Institutes For Health Sciences. (2014). *DILIsym*®. Retrieved from http://www.dilisym.com/Products-Services/dilisym.html (accessed August 29, 2014).

Thienpont B, Tingaud-Sequeira A, Prats E, Barata C, Babin PJ, Raldua D. (2011). Zebrafish eleutheroembryos provide a suitable vertebrate model for screening chemicals that impair thyroid hormone synthesis. *Environmental Science and Technology* 45: 7525-7532. http://www.ncbi.nlm.nih.gov/pubmed/21800831.

Thomas D. (2010). Gene--environment-wide association studies: Emerging approaches. *Nature Reviews Genetics* 11: 259-272. http://www.ncbi.nlm.nih.gov/pubmed/20212493.

Thomas R, Hubbard A, McHale CM, Zhang L, Rappaport SM, Rothman N, et al. (2014). Characterization of changes in gene expression and biochemical pathways at low levels of benzene exposure. *PloS One* 9: e91828. http://www.ncbi.nlm.nih.gov/pubmed/24786086.

Thomas R, McHale CM, Lan Q, Hubbard AE, Zhang L, Vermeulen R, et al. (2013a). Global gene expression response of a population exposed to benzene: A pilot study exploring the use of RNA-sequencing technology. *Environmental and Molecular Mutagenesis* 54: 566-573. http://www.ncbi.nlm.nih.gov/pubmed/23907980.

Thomas R, Phuong J, McHale CM, Zhang L. (2012a). Using bioinformatic approaches to identify pathways targeted by human leukemogens. *International Journal of Environmental Research and Public Health* 9: 2479-2503. http://www.ncbi.nlm.nih.gov/pubmed/22851955.

Thomas RS, Bao W, Chu TM, Bessarabova M, Nikolskaya T, Nikolsky Y, et al. (2009). Use of short-term transcriptional profiles to assess the long-term cancer-related safety of environmental and industrial chemicals. *Toxicological Sciences* 112: 311-321. http://dx.doi.org/10.1093/toxsci/kfp233.

Thomas RS, Black MB, Li L, Healy E, Chu TM, Bao W, et al. (2012b). A comprehensive statistical analysis of predicting *in vivo* hazard using high-throughput *in vitro* screening. *Toxicological Sciences* 128: 398-417. http://www.ncbi.nlm.nih.gov/pubmed/22543276.

Thomas RS, Clewell HJ, 3rd, Allen BC, Yang L, Healy E, Andersen ME. (2012c). Integrating pathway-based transcriptomic data into quantitative chemical risk assessment: A five chemical case study. *Mutation Research* 746: 135-143. http://www.ncbi.nlm.nih.gov/pubmed/22305970.

Thomas RS, Clewell HJ, Allen BC, Wesselkamper SC, Wang NC, Lambert JC, et al. (2011). Application of transcriptional benchmark dose values in quantitative cancer and noncancer risk assessment. *Toxicological Sciences* 120: 194-205. http://dx.doi.org/10.1093/toxsci/kfq355.

Thomas RS, Himmelstein MW, Clewell HJ, 3rd, Yang Y, Healy E, Black MB, et al. (2013b). Cross-species transcriptomic analysis of mouse and rat lung exposed to chloroprene. *Toxicological Sciences* 131: 629-640. http://www.ncbi.nlm.nih.gov/pubmed/23125180.

Thomas RS, Philbert MA, Auerbach SS, Wetmore BA, DeVito MJ, Cote I, et al. (2013c). Incorporating new technologies into toxicity testing and risk assessment: Moving from 21st century vision to a data-driven framework. *Toxicological Sciences*. http://www.ncbi.nlm.nih.gov/pubmed/23958734.

Thomas RS, Pluta L, Yang L, Halsey TA. (2007). Application of genomic biomarkers to predict increased lung tumor incidence in 2-year rodent cancer bioassays. *Toxicological Sciences* 97: 55-64. http://dx.doi.org/10.1093/toxsci/kfm023.

Thomas RS, Wesselkamper SC, Wang NC, Zhao QJ, Petersen DD, Lambert JC, et al. (2013d). Temporal concordance between apical and transcriptional points of departure for chemical risk assessment. *Toxicological Sciences* 134: 180-194. http://www.ncbi.nlm.nih.gov/pubmed/23596260.

Thompson CM, Johns DO, Sonawane B, Barton HA, Hattis D, Tardif R, et al. (2009). Database for physiologically based pharmacokinetic (PBPK) modeling: Physiological data for healthy and health-impaired elderly. *J Toxicol Environ Health B Crit Rev* 12: 1-24. http://www.ncbi.nlm.nih.gov/pubmed/19117207.

Thompson CM, Sonawane B, Barton HA, DeWoskin RS, Lipscomb JC, Schlosser P, et al. (2008). Approaches for applications of physiologically based pharmacokinetic models in risk assessment. *Journal of Toxicology and Environmental Health, Part B* 11: 519-547. http://dx.doi.org/10.1080/10937400701724337.

Tice RR, Austin CP, Kavlock RJ, Bucher JR. (2013). Improving the human hazard characterization of chemicals: A Tox21 update. *Environmental Health Perspectives* 121: 756-765. http://www.ncbi.nlm.nih.gov/pubmed/23603828.

Tietge JE, Degitz SJ, Haselman JT, Butterworth BC, Korte JJ, Kosian PA, et al. (2013). Inhibition of the thyroid hormone pathway in Xenopus laevis by 2-mercaptobenzothiazole. *Aquatic Toxicology* 126: 128-136. http://www.ncbi.nlm.nih.gov/pubmed/23178179.

Timpson NJ, Lindgren CM, Weedon MN, Randall J, Ouwehand WH, Strachan DP, et al. (2009). Adiposity-related heterogeneity in patterns of type 2 diabetes susceptibility observed in genome-wide association data. *Diabetes* 58: 505-510. http://www.ncbi.nlm.nih.gov/pubmed/19056611.

Tokar EJ, Diwan BA, Thomas DJ, Waalkes MP. (2012). Tumors and proliferative lesions in adult offspring after maternal exposure to methylarsonous acid during gestation in CD1 mice. *Archives of Toxicology* 86: 975-982. http://www.ncbi.nlm.nih.gov/pubmed/22398986.

Tokar EJ, Qu W, Waalkes MP. (2011). Arsenic, stem cells, and the developmental basis of adult cancer. *Toxicological Sciences* 120 Suppl 1: S192-203. http://www.ncbi.nlm.nih.gov/pubmed/21071725.

TOPKAT. (2014). *Toxicity prediction by komputer assisted technology, a product of Accelrys, Inc.* Retrieved from http://accelrys.com/mini/toxicology/predictive-functionality.html (accessed August 29, 2014).

Torkamani A, Topol EJ, Schork NJ. (2008). Pathway analysis of seven common diseases assessed by genome-wide association. *Genomics* 92: 265-272. http://www.ncbi.nlm.nih.gov/pubmed/18722519.

U.S. DHHS (U.S. Department of Health and Human Services). (2014). *The Health Consequences of Smoking - 50 Years of Progress*. Rockville, MD: Office of the Surgeon General. Retrieved from http://www.surgeongeneral.gov/library/reports/50-years-of-progress/exec-summary.pdf.

U.S. Surgeon General Centers for Disease Control and Prevention (US); National Center for Chronic Disease Prevention and Health Promotion (US); Office on Smoking and Health (US). (2010). *How Tobacco Smoke Causes Disease: The Biology and Behavioral Basis for Smoking-Attributable Disease: A Report of the Surgeon General*. Retrieved from http://www.ncbi.nlm.nih.gov/books/NBK53017/.

Uehara T, Minowa Y, Morikawa Y, Kondo C, Maruyama T, Kato I, et al. (2011). Prediction model of potential hepatocarcinogenicity of rat hepatocarcinogens using a large-scale toxicogenomics database. *Toxicology and Applied Pharmacology* 255: 297-306. http://dx.doi.org/10.1016/j.taap.2011.07.001.

Vacaru AM, Unlu G, Spitzner M, Mione M, Knapik EW, Sadler KC. (2014). *In vivo* cell biology in zebrafish - providing insights into vertebrate development and disease. *J Cell Sci* 127: 485-495. http://www.ncbi.nlm.nih.gov/pubmed/24481493.

Valcke M, Krishnan K. (2014). Characterization of the human kinetic adjustment factor for the health risk assessment of environmental contaminants. *J Appl Toxicol* 34: 227-240. http://www.ncbi.nlm.nih.gov/pubmed/24038072.

van Leeuwen DM, Pedersen M, Knudsen LE, Bonassi S, Fenech M, Kleinjans JC, et al. (2011). Transcriptomic network analysis of micronuclei-related genes: A case study. *Mutagenesis* 26: 27-32. http://www.ncbi.nlm.nih.gov/pubmed/21164179.

Venkatapathy R, Moudgal CJ, Bruce RM. (2004). Assessment of the oral rat chronic lowest observed adverse effect level model in TOPKAT, a QSAR software package for toxicity prediction. *Journal of Chemical Information and Computer Sciences* 44: 1623-1629. http://www.ncbi.nlm.nih.gov/pubmed/15446819.

Venkatapathy R, Wang NC. (2013). Developmental toxicity prediction. *Methods in Molecular Biology* 930: 305-340. http://www.ncbi.nlm.nih.gov/pubmed/23086848.

Vineis P, van Veldhoven K, Chadeau-Hyam M, Athersuch TJ. (2013). Advancing the application of omics-based biomarkers in environmental epidemiology. *Environ Mol Mutagen* 54: 461-467. http://www.ncbi.nlm.nih.gov/pubmed/23519765.

Visscher PM, Brown MA, McCarthy MI, Yang J. (2012). Five years of GWAS discovery. *Am J Hum Genet* 90: 7-24. http://www.ncbi.nlm.nih.gov/pubmed/22243964.

Vogelstein B, Lane D, Levine AJ. (2000). Surfing the p53 network. *Nature* 408: 307-310. http://www.ncbi.nlm.nih.gov/pubmed/11099028.

Waits ER, Nebert DW. (2011). Genetic architecture of susceptibility to PCB126-induced developmental cardiotoxicity in zebrafish. *Toxicological Sciences* 122: 466-475. http://www.ncbi.nlm.nih.gov/pubmed/21613231.

Walker JD, Carlsen L. (2002). QSARs for identifying and prioritizing substances with persistence and bioconcentration potential. *SAR and QSAR in Environmental Research* 13: 713-725. http://www.ncbi.nlm.nih.gov/pubmed/12570048.

Walker JD, Carlsen L, Hulzebos E, Simon-Hettich B. (2002). Global government applications of analogues, SARs and QSARs to predict aquatic toxicity, chemical or physical properties, environmental fate parameters and health effects of organic chemicals. *SAR and QSAR in Environmental Research* 13: 607-616. http://www.ncbi.nlm.nih.gov/pubmed/12479375.

Wambaugh J, Setzer RW, Reif DM, Gangwal S, Mitchell-Blackwood J, Arnot JA, et al. (2013). High-throughput models for exposure-based chemical prioritization in the ExpoCast project.

Environmental Science and Technology 47: 8479-8488. http://www.ncbi.nlm.nih.gov/pubmed/23758710.

Wambaugh J, Shah I. (2010). Simulating microdosimetry in a virtual hepatic lobule. *Public Library of Science Computational Biology* 6: e1000756. http://www.ncbi.nlm.nih.gov/pubmed/20421935.

Wang I, Zhang B, Yang X, Stepaniants S, Zhang C, Meng Q, et al. (2012a). Systems analysis of eleven rodent disease models reveals an inflammatome signature and key drivers. *Molecular Systems Biology* 8: 594. http://www.ncbi.nlm.nih.gov/pubmed/22806142.

Wang N, Jay Zhao Q, Wesselkamper SC, Lambert JC, Petersen D, Hess-Wilson JK. (2012b). Application of computational toxicological approaches in human health risk assessment I. A tiered surrogate approach. *Regulatory Toxicology and Pharmacology* 63: 10-19. http://www.ncbi.nlm.nih.gov/pubmed/22369873.

Wang N, Rice GE, Teuschler LK, Colman J, Yang RS. (2012c). An *in silico* approach for evaluating a fraction-based, risk assessment method for total petroleum hydrocarbon mixtures. *Journal of Toxicology* 2012: 410143. http://www.ncbi.nlm.nih.gov/pubmed/22496687.

Wang N, Venkatapathy R, Bruce RM, Moudgal C. (2011). Development of quantitative structure-activity relationship (QSAR) models to predict the carcinogenic potency of chemicals. II. Using oral slope factor as a measure of carcinogenic potency. *Regulatory Toxicology and Pharmacology* 59: 215-226. http://www.ncbi.nlm.nih.gov/pubmed/20951756.

Warner CM, Gust KA, Stanley JK, Habib T, Wilbanks MS, Garcia-Reyero N, et al. (2012). A systems toxicology approach to elucidate the mechanisms involved in RDX species-specific sensitivity. *Environmental Science and Technology* 46: 7790-7798. http://www.ncbi.nlm.nih.gov/pubmed/22697906.

Weiss JN, Karma A, MacLellan WR, Deng M, Rau CD, Rees CM, et al. (2012). "Good enough solutions" and the genetics of complex diseases. *Circulation Research* 111: 493-504. http://www.ncbi.nlm.nih.gov/pubmed/22859671.

Wetmore BA, Wambaugh JF, Ferguson SS, Li L, Clewell HJ, 3rd, Judson RS, et al. (2013). Relative impact of incorporating pharmacokinetics on predicting *in vivo* hazard and mode of action from high-throughput *in vitro* toxicity assays. *Toxicological Sciences* 132: 327-346. http://www.ncbi.nlm.nih.gov/pubmed/23358191.

Wetmore BA, Wambaugh JF, Ferguson SS, Sochaski MA, Rotroff DM, Freeman K, et al. (2012). Integration of dosimetry, exposure, and high-throughput screening data in chemical toxicity assessment. *Toxicological Sciences* 125: 157-174. http://www.ncbi.nlm.nih.gov/pubmed/21948869.

White RH, Cote I, Zeise L, Fox M, Dominici F, Burke TA, et al. (2009). State-of-the-science workshop report: Issues and approaches in low-dose-response extrapolation for environmental health risk assessment. *Environ Health Perspect* 117: 283-287. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2649232/.

WHO (World Health Organization). (2010). *Characterization and application of physiologically based pharmacokinetic models in risk assessment*. Geneva, Switzerland. Retrieved from http://www.inchem.org/documents/harmproj/harmproj/harmproj9.pdf.

WHO (World Health Organization). (2012). *State of the Science of Endocrine Disrupting Chemicals*. Retrieved from http://apps.who.int/iris/bitstream/10665/78102/1/WHO_HSE_PHE_IHE_2013.1_eng.pdf (accessed August 29, 2014).

Wignall J, Muratov E, Fourches D, Sedykh A, Tropsha A, Woodruff TJ, et al. (2012). Modeling toxicity values using chemical structure, *in vitro* screening, and *in vivo* toxicity data *51st Annual Meetign of the Society of Toxicology* Poster Abstract # 299. http://www.toxicology.org/AI/Pub/Tox/2012Tox.pdf.

Williams LM, Oleksiak MF. (2011). Ecologically and evolutionarily important SNPs identified in natural populations. *Molecular Biology and Evolution* 28: 1817-1826. http://www.ncbi.nlm.nih.gov/pubmed/21220761.

Wong AP, Rossant J. (2013). Generation of lung epithelium from pluripotent stem cells. *Curr Pathobiol Rep* 1: 137-145. http://www.ncbi.nlm.nih.gov/pubmed/23662247.

Wooding SP, Watkins WS, Bamshad MJ, Dunn DM, Weiss RB, Jorde LB. (2002). DNA sequence variation in a 3.7-kb noncoding sequence 5' of the CYP1A2 gene: Implications for human population history and natural selection. *American Journal of Human Genetics* 71: 528-542. http://www.ncbi.nlm.nih.gov/pubmed/12181774.

Wright FA, Shabalin AA, Rusyn I. (2012). Computational tools for discovery and interpretation of expression quantitative trait loci. *Pharmacogenomics* 13: 343-352. http://www.ncbi.nlm.nih.gov/pubmed/22304583.

Wu W, Doreswamy V, Diaz-Sanchez D, Samet JM, Kesic M, Dailey L, et al. (2011). GSTM1 modulation of IL-8 expression in human bronchial epithelial cells exposed to ozone. *Free Radical Biology and Medicine* 51: 522-529. http://www.ncbi.nlm.nih.gov/pubmed/21621609.

Wu W, Peden D, Diaz-Sanchez D. (2012). Role of GSTM1 in resistance to lung inflammation. *Free Radic Biol Med* 53: 721-729. http://www.ncbi.nlm.nih.gov/pubmed/22683820.

Wyss Institute. (2012). *Wyss Institute to receive up to $37 million from DARPA to integrate multiple organ-on-chip systems to mimic the whole human body*. Retrieved from http://wyss.harvard.edu/viewpressrelease/91/wyss-institute-to-receive-up-to-37-million-from-darpa-to-integrate-multiple-organonchip-systems-to-mimic-the-whole-human-body (accessed August 29, 2014).

Yeung KY, Dombek KM, Lo K, Mittler JE, Zhu J, Schadt EE, et al. (2011). Construction of regulatory networks using expression time-series data of a genotyped population. *Proceedings of the National Academy of Sciences USA* 108: 19436-19441. http://www.ncbi.nlm.nih.gov/pubmed/22084118.

Zaldívar CJM, Wambaugh J, Judson R. (2012). Modelling *in vitro* cell-based assays experiments: cell population dynamics. In *Models of the Ecological Hierarchy*. Amsterdam: Elsevier.

Zeggini E, Weedon MN, Lindgren CM, Frayling TM, Elliott KS, Lango H, et al. (2007). Replication of genome-wide association signals in UK samples reveals risk loci for type 2 diabetes. *Science* 316: 1336-1341. http://www.ncbi.nlm.nih.gov/pubmed/17463249.

Zeise L, Bois FY, Chiu WA, Hattis D, Rusyn I, Guyton KZ. (2012). Addressing human variability in next-generation human health risk assessments of environmental chemicals. *Environmental Health Perspectives* Epub doi: 10.1289/ehp.1205687. http://www.ncbi.nlm.nih.gov/pubmed/23086705.

Zhang L, McHale CM, Rothman N, Li G, Ji Z, Vermeulen R, et al. (2010a). Systems biology of human benzene exposure. *Chemico-Biological Interactions* 184: 86-93. http://www.ncbi.nlm.nih.gov/pubmed/20026094.

Zhang Q, Bhattacharya S, Andersen ME, Conolly RB. (2010b). Computational systems biology and dose-response modeling in relation to new directions in toxicity testing. *Journal of Toxicology and Environmental Health Part B: Critical Reviews* 13: 253-276. http://www.ncbi.nlm.nih.gov/pubmed/20574901.

Zhang Q, Bhattacharya S, Conolly RB, Clewell HJ, 3rd, Kaminski NE, Andersen ME. (in press). Molecular signaling network motifs provide a mechanistic basis for cellular threshold responses. *Environ Health Perspect*. [Epub ahead of print]. http://ehp.niehs.nih.gov/1408244/.

Zhuo W, Zhang L, Zhu B, Qiu Z, Chen Z. (2012). Association between CYP1A1 Ile462Val variation and acute leukemia risk: Meta-analyses including 2164 cases and 4160 controls. *Public Library of Science One* 7: e46974. http://www.ncbi.nlm.nih.gov/pubmed/23056546.

Zoeller RT, Crofton KM. (2005). Mode of action: Developmental thyroid hormone insufficiency--neurological abnormalities resulting from exposure to propylthiouracil. *Critical Reviews in Toxicology* 35: 771-781. http://www.ncbi.nlm.nih.gov/pubmed/16417044.

Zoeller RT, Dowling AL, Herzig CT, Iannacone EA, Gauger KJ, Bansal R. (2002). Thyroid hormone, brain development, and the environment. *Environmental Health Perspectives* 110 Suppl 3: 355-361. http://www.ncbi.nlm.nih.gov/pubmed/12060829.

Zoeller RT, Rovet J. (2004). Timing of thyroid hormone action in the developing brain: Clinical observations and experimental findings. *Journal of Neuroendocrinology* 16: 809-818. http://www.ncbi.nlm.nih.gov/pubmed/15500540.

[page intentionally left blank]

Appendix A
Advancing the Next Generation of Toxicity Testing and Risk Assessment: Government Activities in Europe and the United States

European Union

The European Chemicals Agency (ECHA). In response to environmental concerns, a desire for increased assessment efficiencies, and a desire to reduce reliance on *in vivo* animal testing, the European Union (EU) enacted an expansive new program called Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH) in June 2007. This legislation places greater responsibility on industry to test and manage the risks posed by their chemicals. Under REACH, companies must develop detailed technical dossiers and chemical safety reports and submit them to the European Chemicals Agency (ECHA). About 12,000 chemicals have been registered for consideration with ECHA. Many more chemicals are anticipated in the near future. Additionally, the 7th Amendment to the EU Cosmetics Directive prohibits putting animal tested cosmetics on the market in Europe after 2013. Although current alternative methods more closely resemble traditional methods, the EU has invested 50M Euros in a research program to further next generation methods (OECD 2014c). Current ECHA guidance is available on using quantitative structure activity relationships (QSARs), *in vitro* assays, and read across (also called near analog structure activity relationships) to support assessments.

REACH and the 7th amendment will significantly impact nearly all multinational companies and are important drivers for the development and use of new molecular based methodologies. Europe's chemical trade accounts for about 40% of the global market, involving 27 countries and almost half a billion people.

The Joint Research Centre (JRC) is the scientific and technical arm of the European Commission. It provides scientific advice and technical support to EU policies. The JRC has seven scientific institutes (featuring laboratories and research facilities) located at five sites: Belgium, Germany, Italy, the Netherlands, and Spain. The JRC's Institute for Health and Consumer Protection's main research relevant to NexGen includes integrated risk and benefit assessments of chemical substances; fit for purpose analytical tools to help ensure the safety of food and consumer products; and optimization and validation of methods that reduce the reliance on animal tests in the safety assessment of chemicals.

U.S. Activities

Several documents have guided the NexGen effort, including the Strategic Plan for the Future of Toxicity Testing and Risk Assessment at the U.S. Environmental Protection Agency (EPA 2009b), the Toxicology in the 21st Century (Tox21) strategy, and the National Institutes of Health Strategic Plan (NIEHS 2014e). Ongoing research activities of several federal agencies that have informed and continue to inform the NexGen effort are described below.

The Centers for Disease Control and Prevention (CDC) has several groups involved in systems biology and computational environmental health and occupational research. The ***National Center for Environmental Health (NCEH)*** and ***Agency for Toxic Substances and Disease Registry (ATSDR)*** scientists in the Computational Toxicology Laboratory have applied several new approaches for improving chemical risk assessments. They have mined the National Health and Nutrition Examination Survey (NHANES) data set to obtain high quality analytical and human health information, which is representative of the general U.S. population, and used computer modeling to identify sensitive populations for health outcomes at environmental exposure levels. A second project involved use of NHANES public health genomics data to identify allelic differences in ALA dehydratase for susceptibility to lead induced hypertension. Another concerned the development and application of QSAR, physiologically based pharmacokinetic (PBPK), and molecular docking approaches. These studies involved both data mining of the published scientific literature and collaborative laboratory studies with scientists at the Food and Drug Administration (FDA).

The National Institute for Occupational Safety and Health (NIOSH) is investigating susceptibility gene variants that contribute to the development and severity of occupational diseases using high density and high throughput (HT) genotyping platforms. Understanding the genetic contribution to the development, progression, and outcomes of complex occupational diseases will help improve the accuracy of risk assessment and improve safe exposure levels for genetically susceptible groups in the workforce. ***The FDA National Center for Toxicological Research (NCTR)*** is conducting translational research to develop a scientifically sound basis for regulatory decisions and reduce risks associated with FDA regulated products. NCTR research evaluates biological effects of potentially toxic chemicals, defines the complex mechanisms that govern their toxicity, identifies the critical biological events in the expression of toxicity, discovers

biomarkers, and develops new scientific tools and methods to improve assessment of human exposure, susceptibility, and risk. Examples of tools created by NCTR include ArrayTrack™, Decision Forest, Endocrine Disruptor Knowledge Base (EDKB), Gene Ontology for Functional Analysis (GOFFA), and SNPTrack. Efforts include the MicroArray Quality Control (MAQC) consortia.

The National Institutes of Health (NIH) National Center for Advancing Translational Sciences (NCATS) conducts research to resolve scientific and technical challenges that might cause barriers to the efficient development of new treatments and tests to improve human health. The National Chemical Genomics Center (NCGC) at the National Center for Advancing Translational Sciences applies high throughput screening (HTS) assay guidance, informatics, and chemistry resources for NCAT s Re engineering Translational Sciences research projects. Specifically, NCGC research programs include assay development and HTS, and participation in Tox21. NCGC Assay Biology Teams are researching optimization of biochemical, cellular, and model organism based assays submitted by the biomedical research community for HT small molecule screening. The results of these screens (probes) can be used to further examine protein and cell functions and biological processes relevant to physiology and disease (NIH 2014).

The National Human Genome Research Institute (NHGRI) was established by NIH in 1989 to implement the International Human Genome Project to map the human genome. NHGRI has developed programs for a variety of research projects including Encyclopedia of DNA Elements (ENCODE), Gene Expression Omnibus (GEO), and collaborative projects, including the Comparative Toxicogenomic Database (CTD), HapMap, and Gene. Through the application of these tools, NHGRI hopes to gain a greater understanding of human genetic disease, and develop better methods for the detection, prevention, and treatment of genetic disorders.

The National Institute of Environmental Health Science (NIEHS) and the National Toxicology Program (NTP) have played an integral role in the development and application of HTS data. Current research is focused on developing and validating Tox21 approaches to improve hazard identification, characterization, and risk assessment (Birnbaum 2012; Serafimova et al. 2007). The NTP HTS program has three specific goals: (1) prioritizing substances for in depth toxicological evaluation, (2) identifying mechanisms of action for further investigation (e.g., disease associated pathways), and (3) developing predictive models for *in vivo* biological response (i.e., predictive toxicology). NTP is developing innovative and flexible approaches to data integration, both across research programs and across different data types (e.g., HT, mechanistic, animal studies) (Bucher et al. 2011). These efforts seek to integrate results from new techniques with traditional toxicology data to provide a public health context.

The Engineer Research and Development Center (ERDC), the research organization of the U.S. Army Corps of Engineers, conducts research and development in support of warfighters, military installations, and civil works projects involving water resources and environmental missions. The ERDC Toxicogenomics research cluster focuses on using genomics to develop tools to rapidly assess toxicity of military chemicals in a wide range of animals, identifying gene biomarkers of exposure, understanding the mechanisms by which military chemicals cause toxicity, and extrapolating toxicity effects across multiple species. Capabilities of the team include advanced instrumentation to characterize impacts of chemicals on gene expression with high density gene arrays, DNA sequencing, and real time polymerase chain reaction (RT PCR) assays. ERDC Toxicogenomic projects include development of rapid assays to assess whole genome impacts of munitions related compounds, including gene arrays with short exposure screening in daphnia, rat cells, rat livers, and fish; comparison of genomic and behavioral responses of fathead minnows and zebrafish to chemical exposures; conservation of response to nitroaromatics across species; and support for a toxicogenomic assessment framework to integrate predictive toxicology of munitions related compounds.

Several ***EPA Office of Research and Development (ORD) laboratories and centers*** have been involved in NexGen. EPA's National Center for Environmental Assessment (NCEA) has assumed a leadership and coordination role for the NexGen effort. The National Center for Computational Toxicology (NCCT) is the largest component of EPA s Computational Toxicology Research Program. The Center coordinates computational toxicology research on chemical prioritization and screening, informatics, and systems modeling. NCCT research includes the (1) use of informatics, HTS technologies, and systems biology to develop accurate and flexible computational tools that can screen the thousands of chemicals for potential toxicity; and (2) application of mathematical and advanced computer models to help assess chemical hazards and risks. EPA's National Center for Environmental Research (NCER) supports extramural computational toxicology research. The National Health and Environmental Effects Research Laboratory (NHEERL) conducts toxicological, clinical, and epidemiological research to improve the process of human health risk assessments, including development of biological assays and toxicological assessment methods, predictive pharmacokinetic/pharmacodynamic models, and advanced extrapolation methods.

References

Birnbaum LS. (2012). NIEHS and the future of toxicology. Presentation. *FutureTox Meeting of the Society of Toxicology* October 18-19. Arlington, VA.

Bucher JR, Thayer K, Birnbaum LS. (2011). The Office of Health assessment and translation: A problem-solving resource for the National Toxicology Program. *Environmental Health Perspectives* 119: A196-197. http://www.ncbi.nlm.nih.gov/pubmed/21531652.

EPA (U.S. Environmental Protection Agency). (2009b). *Strategic Plan for the Future of Toxicity Testing and Risk Assessment at the U.S. EPA*. Washington, DC: Office of the Science Advisor, Science Policy Council. Retrieved from http://www.epa.gov/spc/toxicitytesting/docs/toxtest_strategy_032309.pdf (accessed August 29, 2014).

NIH (National Institutes of Health). (2014). *NIH Chemical Genomics Center*. Retrieved from http://www.ncats.nih.gov/research/reengineering/ncgc/ncgc.html (accessed August 29, 2014).

NIEHS (National Institute of Environmental Health Sciences). (2014e). *NIEHS Strategic Plan*. Retrieved from http://niehs.nih.gov/about/strategicplan/index.cfm (accessed August 29, 2014).

OECD (Organization for Economic Cooperation and Development). (2014c). *The OECD QSAR Toolbox. Version 3.0*. Retrieved from http://www.oecd.org/env/chemicalsafetyandbiosafety/assessmentofchemicals/theoecdqsartoolbox.htm (accessed August 29, 2014).

Serafimova R, Todorov M, Nedelcheva D, Pavlov T, Akahori Y, Nakai M, et al. (2007). QSAR and mechanistic interpretation of estrogen receptor binding. *SAR and QSAR in Environmental Research* 18: 389-421. http://www.ncbi.nlm.nih.gov/pubmed/17514577

Appendix B
Science Community and Stakeholder Engagement

This appendix provides more details on outreach to the science community and stakeholders. Outreach is a principle of the NexGen framework. The following discussion of engagement efforts is presented in chronological order.

Expert Workshop

EPA convened a 3-day expert workshop, "Advancing the Next Generation of Risk Assessment: The Prototypes Workshop" on November 1–3, 2010, in Research Triangle Park, North Carolina to discuss the draft framework, draft prototypes, ongoing research, and other project elements. Participants were chosen based on their expertise in traditional and more recent approaches in molecular, computational, and systems biology, particularly as that expertise pertains to the draft prototypes. Individuals from various government agencies and stakeholder categories participated in the workshop. Individual advice, rather than consensus, was sought.

The workshop goals were to (1) explore the best way for developing case studies (termed prototypes) that evaluated and demonstrated how molecular biology information can be used in health assessments; (2) discuss a variety of new data types and methods with potential to characterize data-limited chemicals; (3) consider how this information might augment, extend, or replace traditional data in health assessment; and (4) summarize options for expanded future work and research needs. The workshop report with the agenda and list of participants is available online (EPA 2010).

Stakeholder Involvement
Public Dialogue Conference

EPA sponsored a public dialogue conference on February 15 and 16, 2011, in Washington, DC, "Advancing the Next Generation of Risk Assessment." This conference afforded stakeholders the opportunity to learn about NexGen and provide their thoughts on challenges the program faced and its proposed path forward. Approximately 160 participants, representing 11 stakeholder groups, attended the conference (Figure A-1). A conference report was released (EPA 2011a) and videos of the presentations are also available (EPA 2014e).

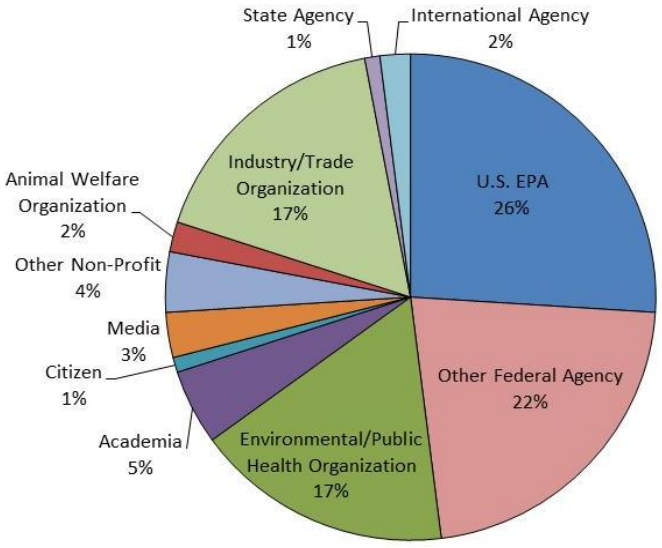


Figure A-1. Categories of Stakeholders that Attended the February 2011 NexGen Public Dialogue Conference (EPA 2011a).

Public-Interest Group Perspectives

Ronald White, a faculty member at Johns Hopkins Bloomberg School of Public Health, conducted informal interviews with several Washington, DC-based representatives of national environmental, public health, and animal welfare public-interest organizations, as part of his research on public engagement. He also developed a Web-based assessment in late 2010 to ascertain, from nongovernmental public-interest organizations, their knowledge and interest in emerging scientific approaches for chemical and pollutant risk assessment. Of the 24 organizations contacted, 8 (33%) responded to the assessment.

A key question in both forums was how relevant the NexGen program is to near-term EPA risk assessment procedures and control policies. The stakeholders generally supported the concept of integrating the results from emerging biological science and analytical techniques into EPA's approach to chemical health-based risk assessment. They also raised concerns regarding the potential to overstate the utility of NexGen approaches; how NexGen prototypes will address key methodological issues; and transparency, meaningful public engagement, and applying the approaches in risk management.

Business Community Perspectives

Dr. Gerald Poje, an environmental health consultant and former member of the U.S. Chemical Safety and Hazard Investigation Board, conducted informal interviews with industry and business representatives. He met with individuals representing the specialty chemical manufacturing and pharmaceutical industries and the retail and energy sectors. The participants generally were optimistic about advances in risk assessment, identifying two potential advantages: (1) better prioritization of the needs for more expensive and longer duration whole-animal testing and (2) saving time and money while rationalizing decisions using high-throughput and other Tiers 1 and 2 data. They suggested that NexGen's success will depend on EPA's ability to prove the value of the tiered approach to EPA's emerging risk assessments, the Agency's investment in the long-term iterative NexGen research effort, and the timely and effective communication of the evidence to support science-based risk assessment. Some in the business community expressed concern over whether EPA could develop the expertise to guide the program to a successful conclusion. Winning over a larger community skeptical of new approaches and the complex associated science might be challenging but such challenges are considered surmountable if EPA can build capacity and communicate how new data types and approaches can be used for risk assessment.

Continued Engagement with the Science Community and the Public

In 2012, the Science Advisory Board and the Board of Scientific Counselors reviewed aspects of the NexGen program as part of their evaluations of EPA's computational toxicology research (BOSC 2010; SAB 2013). Both boards commended EPA's Computational Toxicology Research Program's efforts to advance hazard/risk assessment and made recommendations for its continued success: Continue further research, engage the scientific community and stakeholders, disseminate scientific findings more broadly, gather user feedback from the public, and improve data access.

The National Academies formed the Standing Committee on Use of Emerging Science for Environmental Health Decisions to facilitate communication among government, industry, environmental groups, and academia about scientific advances useful in identifying, quantifying, and controlling environmental impacts on human health. New methods and approaches are explored in workshops, providing a public forum for exchanging information and discussing potential implications for environmental health decisions. These workshops facilitated discussion among the scientific community during the development of the NexGen prototypes.

As mentioned in the introduction, the external peer-review and public comments on the draft NexGen report also have been considered. Changes have been incorporated as appropriate, in this final version.

References

BOSC (Board of Scientific Counselors). (2010). *Review of the EPA Office of Research and Development's (ORD) Computational Toxicology Research Program (CTRP)*. Washington, DC. Retrieved from http://www.epa.gov/osp/bosc/pdf/ctox1004rpt.pdf (accessed August 29, 2014).

EPA (U.S. Environmental Protection Agency). (2010). *Advancing the Next Generation (NexGen) of Risk Assessment: The Prototypes Workshop*. Research Triangle Park, NC. Retrieved from http://www.epa.gov/risk/nexgen/docs/NexGen-Prototypes-Workshop-Summary.pdf (accessed August 29, 2014).

EPA (U.S. Environmental Protection Agency). (2011a). *Advancing the Next Generation (NexGen) of Risk Assessment: Public Dialogue Conference Report*. Washington, DC. Retrieved from http://www.epa.gov/risk/nexgen/docs/NexGen-Public-Conf-Summary.pdf (accessed August 29, 2014).

EPA (U.S. Environmental Protection Agency). (2014e). *Advancing the Next Generation (NexGen) of Risk Assessment: Conferences and Workshops*. Washington, DC. Retrieved from http://www.epa.gov/risk/nexgen/workshops.htm (accessed August 29, 2014).

SAB (Science Advisory Board). (2013). *Draft SAB advice on advancing the application of computational toxicology research for human health risk assessment*. Washington, DC. Retrieved from http://yosemite.epa.gov/sab/sabproduct.nsf/46963ceebabd621905256cae0053d5c6/F3315D0EE2EDC11285257B0A005F5B7E/$File/CompTox-edited+1-29-13.pdf (accessed August 29, 2014).

Appendix C
Principles and Methods for Uncertainty and Variability Analysis

In *A Risk Characterization Framework for Decision-Making at the Food and Drug Administration*, the National Research Council (NRC) noted methods for uncertainty analyses. These methods are applicable to new and traditional data types.

A white paper written for EPA outlines a general hierarchy of methods that can be used to estimate quantities when uncertainty about their "true" values is substantial (Frey et al. 2003). Four general categories of methods are described without any implied preferences or priorities:

1. statistical methods based on empirical data, which use classical statistics to draw inferences from "hard" data alone;
2. statistical methods based on judgment, in which expert judgments and Bayesian approaches to statistical analysis are included, often in combination with "hard" data;
3. other quantitative methods that involve approaches not based on probability theory, such as interval methods, fuzzy methods, and meta-analytic methods; and
4. qualitative methods that can be used when key aspects of uncertainty cannot be captured by quantitative methods.

In the report *Science and Decisions*, NRC articulated principles for uncertainty and variability analyses. "The principles in Box 4-7 are consistent with and expand on the 'Principles for Risk Analysis' originally established in 1995, noted as useful by the National Research Council (NRC 2007b), and recently re-released by the Office of Management and Budget and the Office of Science and Technology Policy (OMB/OSTP 2007)" (NRC 2009). In another report, *Environmental Decisions in the Face of Uncertainty* (NRC 2013), NRC recommended the following principles for uncertainty and variability analysis:

- Risk assessments should provide a quantitative, or at least qualitative, description of uncertainty and variability consistent with available data. The information required to conduct detailed uncertainty analyses might not be available in many situations.
- In addition to characterizing the full population at risk, attention should be directed to vulnerable individuals and subpopulations that might be particularly susceptible or more highly exposed.
- The depth, extent, and detail of the uncertainty and variability analyses should be commensurate with the importance and nature of the decision to be informed by the risk assessment and with what is valued in a decision. This might best be achieved by engaging assessors, managers, and stakeholders early in the nature and objectives of the risk assessment and terms of reference (which must be clearly defined).
- The risk assessment should compile or otherwise characterize the types, sources, extent, and magnitude of variability and substantial uncertainties associated with the assessment. To the extent feasible, treatment of uncertainties among the different components of a risk assessment and among different policy options being compared should be homologous.

- To maximize public understanding of, and participation in, risk-related decision-making, a risk assessment should explain the basis and results of the uncertainty analysis with sufficient clarity to be understood by the public and decision-makers. The uncertainty assessment should not be a significant source of delay in releasing an assessment.
- Uncertainty and variability should be kept separate conceptually in the risk characterization.

References

Frey C, Crawford-Brown D, Zheng J, Loughlin D USE OAQPS. (2003). *Hierarchy of methods to characterize uncertainty: State of science of methods for describing and quantifying uncertainty*. RTP, NC. Retrieved from http://www4.ncsu.edu/~frey/reports/Frey_CrawfordBrown_Zheng_Loughlin_2003.pdf.

NRC (National Research Council). (2007b). *Toxicity Testing in the 21st Century: A Vision and a Strategy*. Washington DC. Retrieved from http://dels.nas.edu/resources/static-assets/materials-based-on-reports/reports-in-brief/Toxicity_Testing_final.pdf (accessed August 29, 2014).

NRC (National Research Council). (2009). *Science and Decisions: Advancing Risk Assessment*. Washington, DC. Retrieved from http://www.nap.edu/catalog/12209.html (accessed August 29, 2014).

NRC (National Research Council). (2013). *Environmental Decisions in the Face of Uncertainty*. (9780309130349). The National Academies Press. Retrieved from http://www.nap.edu/openbook.php?record_id=12568.

Appendix D
Glossary

Glossary Term	Description
AC_{50}	The concentration at which activity is 50 percent of its maximum. This value is useful in comparing assay results.
adverse outcome pathway (AOP); AOP network	An AOP analytical construct that describes a sequential chain of causally linked events at different levels of biological organization that lead to an adverse health or ecotoxicological effect. AOPs are the central element of a toxicological knowledge framework being built to support chemical risk assessment based on mechanistic reasoning. AOP networks are the interrelated set of AOPs that generally underlie disease and are generally analogous to National Institutes of Health's National Center for Biotechnology Information Diagrams for specific diseases. Organization for Economic Cooperation and Development (OECD). The OECD Adverse Outcome Pathway (AOP) program. Retrieved from http://www.oecd.org/env/ehs/testing/adverse-outcome-pathways-molecular-screening-and-toxicogenomics.htm (accessed August 29, 2014). Ankley GT; Bennett RS; Erickson RJ; Hoff DJ; Hornung MW; Johnson RD; Mount DR; Nichols JW Russom CL; Schmieder PK; Serrrano JA; Tietge JE; Villeneuve DL. (2010). Adverse outcome pathways: A conceptual framework to support ecotoxicology research and risk assessment. Environmental Toxicology and Chemistry 29 (3): 730-741. http://service004.hpc.ncsu.edu/toxicology/websites/journalclub/linked_files/Fall10/Environ%20Toxicol%20Chem%202010%20Ankley.pdf.
ArrayTrack™	Publicly available toxicogenomics software for DNA microarrays. It contains three integrated components: (1) a database (MicroarrayDB) that stores microarray data and associated toxicological information; (2) tools (TOOL) for data visualization and analysis; and (3) libraries (LIB) that provide curated functional data from public databases for data interpretation. Using ArrayTrack™, an analysis method can be selected from TOOL and applied to selected microarray data stored in the MicroarrayDB. Analysis results can be linked directly to pathways, gene ontology, and other functional information stored in LIB. Food and Drug Administration. ArrayTrack™ FAQs. Available online at http://www.fda.gov/ScienceResearch/BioinformaticsTools/Arraytrack/ucm135070.htm (accessed August 29, 2014).
assay	1. The process of quantitative or qualitative analysis of a component of a sample; or 2. Results of a quantitative or qualitative analysis of a component of a sample. National Library of Medicine. IUPAC Glossary of Terms Used in Toxicology, 2nd Ed. Available online at http://sis.nlm.nih.gov/enviro/iupacglossary/frontmatter.html (accessed August 29, 2014).
benchmark dose (BMD)	An approach that uses dose-response modeling used to help describe dose-response relationships, that is, the percent of the population exhibiting an adverse effect(s) associated with specific doses of a chemical. The BMD corresponds to specific response levels near the low end of the observable range of the data. The BMD lower limit (BMDL) is a statistical lower confidence limit on the dose at the BMD

Glossary Term	Description
	U.S. Environmental Protection Agency (2012). Benchmark dose technical guidance. Available online at http://www.epa.gov/raf/publications/pdfs/benchmark_dose_guidance.pdf (accessed August 29, 2014).
bioinformatics	A field of biology in which complex multivariable data from high-throughput screening and genomic assays are interpreted in relation to target identification and effects of sustained perturbations on organs and tissues to make biological discoveries or predictions. This field encompasses all computational methods and theories applicable to molecular biology and areas of computer-based techniques for solving biological problems, including manipulation of models and data sets. National Institutes of Health's (NIH) National Center for Biotechnology Information. Bioinformatics. Available online at http://www.ncbi.nlm.nih.gov/mesh?term=bioinformatics (accessed August 29, 2014).
bioassay	A method of measuring the effects of a biologically active substance using an intermediate *in vivo* or *in vitro* tissue or cell model under controlled conditions. It includes virulence studies in animal fetuses *in utero*, mouse convulsion bioassay of insulin, quantitation of tumor-initiator systems in mouse skin, calculation of potentiating effects of a hormonal factor in an isolated strip of contracting stomach muscle, etc. NIH's National Center for Biotechnology Information. Available online at http://www.ncbi.nlm.nih.gov/mesh?term=bioassay (accessed August 29, 2014).
biomarkers	Measurable and quantifiable biological parameters (e.g., specific enzyme concentrations, specific hormone concentrations, a specific gene phenotype distribution in a population, presence of biological substances) that serve as indices for health- and physiology-related assessments, such as disease risk and environmental exposures. NIH's National Center for Biotechnology Information. Biological Markers. Available online at http://www.ncbi.nlm.nih.gov/mesh?term=biological%20markers (accessed August 29, 2014).
BioSystems Database	A biosystem, or biological system, is a group of molecules that interact in a biological system. One type of biosystem is a biological pathway, which can consist of interacting genes, proteins, and small molecules. Another type of biosystem is a disease, which can involve components such as genes, biomarkers, and drugs. A number of databases provide diagrams showing the components and products of biological pathways along with corresponding annotations and links to literature. The NCBI BioSystems Database was developed as a complementary project to (1) serve as a centralized repository of data; (2) connect the biosystem records with associated literature, molecular, and chemical data throughout the Entrez system; and (3) facilitate computation on biosystems data. NIH's National Center for Biotechnology Information. Available online at http://www.ncbi.nlm.nih.gov/Structure/biosystems/docs/biosystems_about.html (accessed August 29, 2014).

Glossary Term	Description
Comparative Toxicogenomic Database (CTD)™	A publicly available toxicogenomic database on the National Library of Medicine's (NLM) Toxicology Data Network (TOXNET®). The CTD™ elucidates molecular mechanisms by which environmental chemicals affect human disease. It contains manually curated data describing cross-species chemical-gene/protein interactions and chemical- and gene-disease relationships. The results provide insight into the molecular mechanisms underlying variable susceptibility and environmentally influenced diseases. These data also will provide insights into complex chemical-gene and protein interaction networks. National Library of Medicine (2012). Fact Sheet. Comparative Toxicogenomics Database (CTD)™. Available online at http://www.nlm.nih.gov/pubs/factsheets/ctdfs.html (accessed August 29, 2014).
computational models	Computerized predictive tools. Sometimes referred to as "*in silico*" models. U.S. Environmental Protection Agency. Glossary of Terms: Methods of Toxicity Testing and Risk Assessment. Available online at http://www.epa.gov/opp00001/science/comptox-glossary.html (accessed August 29, 2014).
decision context	Decision context seeks to understand and describe what management decisions are being made, why these decisions are made, and the relationship of these decisions to previous and anticipated decisions. For example, decision context tries to answer some of the following questions: Are risks being ranked; if so, why? How will risk information be used in future decisions? Is a change in policy or management under consideration; and if so, what is driving the change and what are the underlying policy objectives? What is the general scope of alternatives under consideration and why? Decision context defines the roles and responsibilities of the ultimate decision maker, stakeholders, and key technical experts in relation to the decision process. Decision context also identifies the constraints within which a decision must be made and outputs that will result from the decision. Structured Decision Making (SDM). (2008). Steps in the Decision Process: Introduction. Available online at http://www.structureddecisionmaking.org/steps/decisioncontext/ (accessed August 29, 2014).
dbSNP	dbSNP is world's largest database for nucleotide variations, and is part of the National Center for Biotechnology Information (NCBI), an internationally respected resource for molecular biology information. As of this date, dbSNP comprises a large cluster of species-specific databases that contain over 12 million nonredundant sequence variations (single nucleotide polymorphisms, insertion/deletions, and short tandem repeats) and over 1 billion individual genotypes from HapMap and other large-scale genotyping activities—more than 200GB of data and growing daily. National Library of Medicine. General Information about dbSNP as a Database Resource. Available online at http://www.ncbi.nlm.nih.gov/books/NBK44469/#Info.what_is_dbsnp (accessed August 29, 2014).

Glossary Term	Description
epigenetics	An emerging field of science that studies heritable changes caused by the activation and deactivation of genes with no change in the underlying DNA sequence of the organism. The word is Greek in origin and literally means over and above (epi) the genome. NIH's National Human Genome Research Institute. Talking Glossary of Genetic Terms. Available online at http://www.genome.gov/glossary/index.cfm?id=528&textonly=true (accessed August 29, 2014).
functional genomics	The study of dynamic cellular processes such as gene transcription, translation, and gene product interactions that define an organism. NIH. Genomics and Advanced Technologies. Available online at http://www.niaid.nih.gov/topics/pathogengenomics/Pages/definitions.aspx (accessed August 29, 2014).
gene-environment interaction	The combined effects of genotypes and environmental factors on phenotypic characteristics. NIH's National Center for Biotechnology Information. Gene-Environment Interaction. Available online at http://www.ncbi.nlm.nih.gov/mesh?term=gene%20environment%20interaction (accessed August 29, 2014).
gene expression	The phenotypic manifestation of a gene or genes by the processes of genetic transcription and genetic translation. NIH's National Center for Biotechnology Information. Gene Expression. Available online at http://www.ncbi.nlm.nih.gov/mesh/68015870 (accessed August 29, 2014).
Gene Expression Omnibus (GEO)	A public repository that archives and freely distributes microarray, next-generation sequencing, and other forms of high-throughput functional genomic data submitted by the scientific community. In addition to data storage, a collection of Web-based interfaces and applications is available to help users query and download the studies and gene expression patterns stored in GEO. NIH's National Center for Biotechnology Information. Gene Expression Omnibus. Frequently Asked Questions. Available online at http://www.ncbi.nlm.nih.gov/geo/info/faq.html (accessed August 29, 2014).
Gene Ontology (GO) database	A product of the Gene Ontology (GO) project. The GO project provides structured, controlled vocabularies and classifications that cover several domains of molecular and cellular biology and are freely available for community use in the annotation of genes, gene products, and sequences. Many model organism databases and genome annotation groups use the GO database and contribute their annotation sets to the GO resource. The GO database integrates the vocabularies and contributed annotations and provides full access to this information in several formats. Members of the GO Consortium continuously work collectively, involving outside experts as needed, to expand and update the GO vocabularies. The GO Web resource also provides access to extensive documentation about the GO project and links to applications that use GO data for functional analyses. Gene Ontology Consortium. The Gene Ontology (GO) database and informatics resource. Nucleic Acids Research 32: Database issue D258-261.

Glossary Term	Description
genetics	The branch of science concerned with the means and consequences of transmission and generation of the components of biological inheritance. Used for mechanisms of heredity and the genetics of organisms, for the genetic basis of normal and pathological states, and for the genetic aspects of endogenous chemicals. It includes biochemical and molecular influence on genetic material. NIH's National Center for Biotechnology Information. Genetics. Available online at http://www.ncbi.nlm.nih.gov/mesh/68005823 (accessed August 29, 2014).
genome-wide association study (GWAS)	An approach used in genetics research to associate specific genetic variations with particular diseases. The method involves scanning the genomes from many different people and looking for genetic markers that can be used to predict the presence of a disease. Once such genetic markers are identified, they can be used to understand how genes contribute to the disease and develop better prevention and treatment strategies. NIH. Talking Glossary of Genetic Terms: Genome-wide Association Studies (GWAS). National Human Genome Research Institute. Available online at http://www.genome.gov/glossary/index.cfm?id=91&textonly=true (accessed August 29, 2014).
green chemistry	The design of chemical products and processes to reduce or eliminate the use and generation of hazardous substances. Green Chemistry framework includes three main principles: (1) to incorporate sustainable designs across all stages of the chemical lifecycle, (2) to reduce the hazard of chemical products and processes by design, and (3) to work as a cohesive set of design criteria. Twelve design criteria have been developed to fulfill these three principles (prevention, atom economy, less hazardous chemical synthesis, designing safer chemicals, safer solvents and auxiliaries, design for energy efficiency, use of renewable feedstocks, reduce derivatives, catalysis, design for degradation, real-time analysis for pollution prevention, and inherently safer chemistry for accident prevention). Anastas, P, Eghbali, N. (2010). Green chemistry: Principles and practice. Chem Soc Rev 39 (1): 301-312.
high-content screening (HCS) assay	A method with multiple simultaneous readouts used to analyze system dynamics at any specified level of organization, but generally referring to the whole body, whole cell, or subcellular level of organization. Assay development guidelines for image-based high content screening, high content analysis and high content imaging. William Buchser, Ph.D., Mark Collins, Ph.D., Tina Garyantes, Ph.D., Rajarshi Guha, Ph.D., Steven Haney, Ph.D., Vance Lemmon, Ph.D., Zhuyin Li, Ph.D., and O. Joseph Trask, Jr, B.S. Available online at http://www.ncbi.nlm.nih.gov/books/NBK100913/ (accessed August 29, 2014).
high-throughput screening (HTS) assay	A rapid method of measuring the effect of an agent in a biological or chemical assay. The assay usually involves some form of automation or a way to conduct multiple assays at the same time using sample arrays. NIH's National Center for Biotechnology Information. High-Throughput Screening Assays. Available online at http://www.ncbi.nlm.nih.gov/mesh?term=high%20throughput%20screening%20method (accessed August 29, 2014).

Glossary Term	Description
human toxome	The entirety of pathways of toxicity in humans. A project sponsored by an NIEHS grant (R01ES020750) is an initiative to map the human toxome using systems toxicology approaches. The Human Toxome Project. Available online at http://humantoxome.com/ (accessed August 29, 2014).
in silico	See "computational models" above. National Library of Medicine. (2012). IUPAC Glossary of Terms Used in Toxicology, 2nd Ed. Available online at http://sis.nlm.nih.gov/enviro/iupacglossary/frontmatter.html (accessed August 29, 2014).
IVIV extrapolation (IVIVE)	A method that uses determinations of protein binding, liver/kidney clearance, and oral uptake to estimate ranges of oral human exposures leading to tissue/plasma concentrations similar to *in vitro* point-of-departure concentrations.
knowledgebase	Provides an alternative approach for storing and searching the complete networks of highly interconnected information produced by linking bioassays and pathways. Developed decades ago to codify human knowledge so that they could be used efficiently to support decisions, knowledgebases are finding practical applications in meaningfully organizing vast amounts of linked biological data using ontologies.
knowledge mining	Knowledge mining is the computerized extraction of useful, often previously unknown, information from large databases or data sets using sophisticated data search capabilities and statistical algorithms to discover patterns and correlations and then to interpret this new information in the context of systems biology to create new knowledge.
Kyoto Encyclopedia of Genes and Genomes (KEGG)	A database resource that integrates genomic, chemical, and systemic functional information. In particular, gene catalogs from completely sequenced genomes are linked to higher level systemic functions of the cell, the organism, and the ecosystem. KEGG is a reference knowledgebase for integration and interpretation of large-scale data sets generated by genome sequencing and other high-throughput experimental technologies. Kanehisa Laboratories. KEGG: Kyoto encyclopedia of genes and genomes. Available online at http://www.genome.jp/kegg/ (accessed August 29, 2014).
lift	Lift is a measure of how much better prediction results are using a model than could be obtained by chance. For example, say 2 percent of customers who receive a catalog in the mail make a purchase, and when a model is used to select catalog recipients, 10 percent make a purchase. The lift for the model would be 10/2 or 5. Oracle. Glossary: "Lift." Available online at http://docs.oracle.com/cd/B28359_01/datamine.111/b28129/glossary.htm (accessed August 29, 2014).
mechanism of action	A "sequence of key events and processes, starting with interaction of an agent with a cell, proceeding through operational and anatomical changes, and resulting in an adverse health effect"

Glossary Term	Description
	OECD (Organization for Economic Cooperation and Development). (2013a). Guidance document on developing and assessing adverse outcome pathways. Retrieved from http://search.oecd.org/officialdocuments/displaydocumentpdf/?cote=env/jm/mono%282013%296&doclanguage=en
	OECD (Organization for Economic Cooperation and Development). (2014c). Other activities on molecular screening and toxicogenomics. Retrieved from http://www.oecd.org/env/ehs/testing/toxicogenomics.htm
meta-analysis	A quantitative, formal, epidemiological study design used to assess previous research studies systematically to derive conclusions about that body of research. Outcomes from a meta-analysis can include a more precise estimate of the effect of treatment or risk factor for disease, or other outcomes, than any individual study contributing to the pooled analysis.
	Ramasamy A, Mondry A, Holmes CC, Altman DG. (2008). Key issues in conducting a meta-analysis of gene expression microarray datasets. Public Library of Science Medicine 5: e184.
metabolomics	Type of global molecular analysis that involves identifying and quantifying the metabolome—all metabolites present in a cell at a given time.
	Department of Energy. Human Genome Project Information: Genome Glossary. Available online at http://web.ornl.gov/sci/techresources/Human_Genome/glossary.shtml (accessed August 29, 2014).
microarray analysis	The simultaneous analysis, on a microchip, of multiple samples or targets arranged in an array format.
	NIH's National Center for Biotechnology Information. Microarray Analysis. Available online at http://www.ncbi.nlm.nih.gov/mesh/?term=microarray%20analysis (accessed August 29, 2014).
microarray technology	A technology used to study the expression of many genes at once. It involves placing thousands of gene sequences in known locations on a glass slide called a gene chip. A sample containing DNA or RNA is placed in contact with the gene chip. Complementary base pairing between the sample and the gene sequences on the chip produces light that is measured. Areas on the chip producing light identify genes that are expressed in the sample.
	NIH's National Human Genome Research Institute. Talking Glossary of Genetic Terms. Available online at http://www.genome.gov/glossary/index.cfm?id=125&textonly=true (accessed August 29, 2014).
mode of action (MOA)	The key steps in the toxic response after chemical interaction at the target site that is responsible for the physiological outcome or pathology of the chemical; how chemicals perturb normal biological function.
	U.S. Environmental Protection Agency. Glossary of Terms: Methods of Toxicity Testing and Risk Assessment. Available online at http://www.epa.gov/opp00001/science/comptox-glossary.html (accessed August 29, 2014).

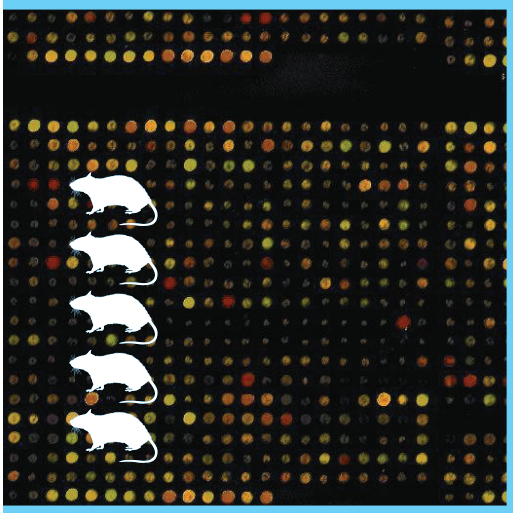
Glossary Term	Description
molecular biology	The branch of biology that deals with the molecular basis of biological activity based on knowledge from biology and chemistry with a focus on genetics and biochemistry.
molecular epidemiology	The use of all types of biological markers in the investigation of the cause, distribution, prevention, and treatment of disease, in which biological markers are used to represent exposures, intervening factors, susceptibility, intermediate pathological events, preclinical and clinical disease for prognosis. Schulte PA, Rothman N, Hainaut FP, Smith MT, Boffetta P, Perea FP. (2011). Molecular epidemiology: linking molecular scale insights into population impacts. In: N Rothman, P Hainaut, P Schulte, M Smith, P Boffetta, F Perea (eds.). Molecular epidemiology: principles and practices. IARC Sci Publ. 2011;(163):1-7.
omics	Refers to a broad field of study in biology, ending in the suffix ''-omics'' such as genomics, proteomics, transcriptomics. U.S. Environmental Protection Agency. Glossary of Terms: Methods of Toxicity Testing and Risk Assessment. Available online at http://www.epa.gov/opp00001/science/comptox-glossary.html (accessed August 29, 2014).
ontology	Defines types of data (e.g., chemicals, genes, assays, interactions, pathways, cells, species) and their interrelationships (chemicals “activate” proteins; assays “measure” changes in proteins; genes are “part of” pathways, etc.).
physiologically based pharmacokinetic (PBPK)	PBPK models emulate pharmacokinetics in the body and are used to estimate the dose to a target tissue or organ by accounting for the rates of absorption, distribution among target organs and tissues, metabolism, and excretion. PBPK models also are often referred to as physiologically based toxicokinetic (PBTK) models in risk assessment to clearly distinguish the chemical as a toxicant. Both terms are in common use, and might appear in the text of this document. They relate to the same kind of model and are interchangeable. EPA. (2014g). Vocabulary Catalog List Detail - Integrated Risk Information System (IRIS) Glossary August 31, 2011. Retrieved from http://ofmpub.epa.gov/sor_internet/registry/termreg/searchandretrieve/glossariesandkeywordlists/search.do?details=&glossaryName=IRIS%20Glossary
pharmacokinetics	Pharmacokinetics has complex meaning that encompasses both a remedy and a toxicant (and more broadly any biologically active substance); risk assessors sometimes use the word “toxicokinetics” to distinguish the chemical as a toxicant. Both terms are in common use, and might appear in the text of this document. They relate to the same processes and are interchangeable.
phenotype	An individual's observable traits, such as height, eye color, and blood type. The genetic contribution to the phenotype is called the genotype. Some traits are largely determined by the genotype, while other traits are largely determined by environmental factors. National Human Genome Research Institute. Talking Glossary of Genetic Terms. Available online at http://www.genome.gov/glossary/index.cfm?id=152&textonly=true (accessed August 29, 2014).

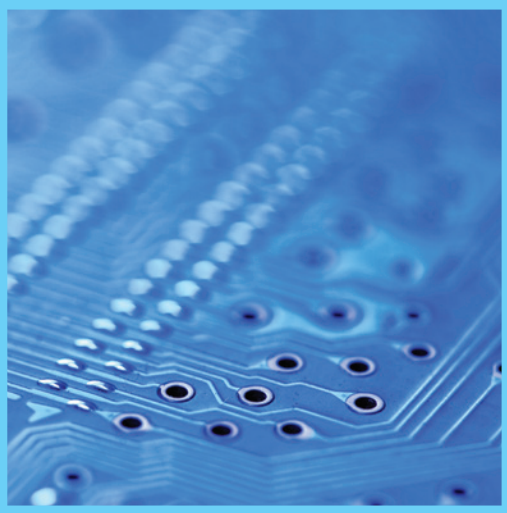
Glossary Term	Description
polymerase chain reaction (PCR)	A method for amplifying a DNA base sequence using a heat-stable polymerase and two 20-base primers, one complementary to the (+) strand at one end of the sequence to be amplified and one complementary to the (-) strand at the other end. Because the newly synthesized DNA strands can subsequently serve as additional templates for the same primer sequences, successive rounds of primer annealing, strand elongation, and dissociation produce rapid and highly specific amplification of the desired sequence. PCR also can be used to detect the existence of the defined sequence in a DNA sample. Department of Energy. Human Genome Project Information: Genome Glossary. Available online at http://web.ornl.gov/sci/techresources/Human_Genome/glossary.shtml (accessed August 29, 2014).
probe	Single-stranded DNA or RNA molecules of specific base sequence, labeled either radioactively or immunologically, that are used to detect the complementary base sequence by hybridization. Department of Energy. Human Genome Project Information: Genome Glossary. Available online at http://web.ornl.gov/sci/techresources/Human_Genome/glossary.shtml#P (accessed August 29, 2014).
proteomics	The study of the function of all expressed proteins. U.S. Environmental Protection Agency (2012). Glossary of Terms: Methods of Toxicity Testing and Risk Assessment. Available online at http://www.epa.gov/opp00001/science/comptox-glossary.html (accessed August 29, 2014).
quantitative structure activity relationship (QSAR)	A mathematical relationship between a quantifiable aspect of chemical structure and a chemical property or reactivity or a well-defined biological activity, such as toxicity. Using a sample set of chemicals, a relationship is established between one or many physical-chemical properties a chemical possesses due to its structure and a chemical property or biological activity of concern. This mathematical expression is then used to predict the chemical property or biological response expected from other chemicals with similar structures. It is based on the presumption that similar molecules or chemical structures have similar properties or biological activities or toxicity potential. U.S. Environmental Protection Agency. Glossary of Terms: Methods of Toxicity Testing and Risk Assessment. Available online at http://www.epa.gov/opp00001/science/comptox-glossary.html (accessed August 29, 2014).
QSAR Toolbox	A software application intended for use by government, the chemical industry, and other stakeholders in filling gaps in (eco)toxicity data needed for assessing the hazards of chemicals. The Toolbox incorporates information and tools from various sources into a logical workflow. Crucial to this workflow is grouping chemicals into chemical categories. The seminal features of the Toolbox are identification of relevant structural characteristics and the potential mechanism or mode of action of a target chemical, identification of other chemicals that have the same structural characteristics or mechanism/mode of action (or both), and use of existing experimental data to fill the data gap(s).

Glossary Term	Description
	QSAR Toolbox. About: What does the QSAR Toolbox do? Available online at http://www.qsartoolbox.org/ (accessed August 29, 2014).
Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH)	A regulation of the European Union, adopted to improve the protection of human health and the environment from the risks that can be posed by chemicals, while enhancing the competitiveness of the European Union chemicals industry. It also promotes alternative methods for the hazard assessment of substances to reduce the number of tests on animals. REACH requirements went into effect on 1 June 2007 and are implemented by the European Chemicals Agency (ECHA). European Chemicals Agency. About us. Available online at http://echa.europa.eu/about-us (accessed August 29, 2014).
reference value	A generic term for an estimate of an exposure for a given duration to the human population (including susceptible subgroups) that is likely to be without an appreciable risk of adverse health effects over a lifetime. Examples of numerical reference values include the reference dose (RfD) and reference concentration (RfC). U.S. Environmental Protection Agency's Integrated Risk Information System (IRIS) Glossary. (2012). Vocabulary Catalog List Detail. Available online at http://ofmpub.epa.gov/sor_internet/registry/termreg/searchandretrieve/glossariesandkeywordlists/search.do?details=&glossaryName=IRIS%20Glossary (accessed August 29, 2014).
reverse toxicokinetics (RTK)	Also known as reverse dosimetry, refers to the use of a pharmacokinetic model to estimate external dose (exposure) from a known internal concentration. The method uses a one-compartment model and makes default assumptions such as chemicals are eliminated wholly through metabolism and renal excretion; renal excretion is a function of the glomerular filtration rate and the fraction of unbound chemical in the blood (i.e., no active transport); and oral absorption is 100 percent. Using these assumptions, the plasma concentration of the chemical at steady state per unit dose then can be estimated. The two experimental chemical-specific parameters required to generate an estimate are the rate of disappearance of parent via hepatic metabolism (intrinsic clearance) and fraction bound (or conversely unbound) to plasma proteins. Both parameters can be measured experimentally in a relatively high-throughput manner. Judson RS; Kavlock RJ; Setzer RW; Hubal EA; Martin MT; Knudsen TB; Houck KA; Thomas RS; Wetmore BA; Dix DJ. (2011). Estimating toxicity-related biological pathway altering doses for high-throughput chemical risk assessment. Chem Res Toxicol Chem Res Toxicol 24 (4): 451-462.
rule	A rule describes an association between elements on the left-hand side of the rule and items on the right-hand side of the rule. For instance, the rule [diapers, cola] => [milk] in a supermarket database might mean that when customers bought diapers and cola, they also purchased milk.
SNPs	Refers to single nucleotide polymorphisms, which are single nucleotide variations in a genetic sequence that occur at appreciable frequency in the population. NIH's National Center for Biotechnology Information. SNPs. Available online at http://www.ncbi.nlm.nih.gov/mesh?term=SNPS (accessed August 29, 2014).

Glossary Term	Description
systems biology	A scientific approach that combines the principles of engineering, mathematics, physics, and computer science with extensive experimental data to develop a quantitative as well as a deep conceptual understanding of biological phenomena, permitting prediction and accurate simulation of complex (emergent) biological behaviors. Wanjek, C. (2011). Systems biology as defined by NIH. The NIH Catalyst 19 (6): November-December. http://irp.nih.gov/catalyst/v19i6/systems-biology-as-defined-by-nih.
toxicokinetics	Risk assessors will sometimes use the word toxicokinetics to distinguish the chemical as a toxicant from a drug and the more traditional use of the word pharmacokinetics. Both terms are in common use, and appear in the text. They relate to the same processes, and are interchangeable.
Tox21	Tox21 is a collaborative effort among four U.S. government agencies (U.S EPA, NIEHS/NTP, NCATs, U.S. FDA) to develop more efficient approaches to predict how chemicals might affect human health. In Tox21 studies, substances are tested using *in vitro* rodent and human cell-based and biochemical assays and lower organisms as model systems. These assays are run at higher throughput and lower cost than animal tests; in some cases, many thousands of chemicals can be tested in a few days. Data from these assays can potentially be used to prioritize substances for further evaluation, inform our understanding of mechanisms of action, and develop improved predictive models for toxicity. Ultimately, test approaches developed and data collected via the Tox21 initiative could enable agencies to reduce their reliance on animal data for establishing regulations for safe handling of chemicals. ICCVAM will evaluate testing approaches developed through the Tox21 collaboration that show promise for regulatory applications and make recommendations on their use to federal agencies. National Toxicology Program. Available online http://ntp.niehs.nih.gov/iccvam/docs/annrpt/iccvam-bienrpt-2014-508.pdf (accessed August 29, 2014).
ToxCast	A major part of EPA's CompTox research is the Toxicity Forecaster (ToxCast™). ToxCast is a multiyear effort launched in 2007 that uses automated chemical screening technologies (called "high-throughput screening assays") to expose living cells or isolated proteins to chemicals. The cells or proteins are then screened for changes in biological activity that suggest potential toxic effects and eventually potential adverse health effects. These innovative methods have the potential to limit the number of required laboratory animal-based toxicity tests while quickly and efficiently screening large numbers of chemicals. U.S. Environmental Protection Agency. Available online http://www.epa.gov/ncct/toxcast (assessed August 29, 2014)
toxicogenomics	Study of the roles that genes play in the biological responses to environmental toxicants and stressors by the collection, interpretation, and storage of information about gene and protein activity. U.S. Environmental Protection Agency. Glossary of Terms: Methods of Toxicity Testing and Risk Assessment. Available online at http://www.epa.gov/opp00001/science/comptox-glossary.html (accessed August 29, 2014).

Glossary Term	Description
transcription	The biosynthesis of RNA carried out on a template of DNA. The biosynthesis of DNA from an RNA template is called reverse transcription. NIH's National Center for Biotechnology Information. Transcription. Available online at http://www.ncbi.nlm.nih.gov/mesh/68014158 (accessed August 29, 2014).
transcriptome	The pattern of gene expression, at the level of genetic transcription, in a specific organism or under specific circumstances in specific cells. NIH's National Center for Biotechnology Information. Transcriptome. Available online at http://www.ncbi.nlm.nih.gov/mesh/68059467 (accessed August 29, 2014).
transcriptomics	The study of gene expression at the RNA level. U.S. Environmental Protection Agency. Glossary of Terms: Methods of Toxicity Testing and Risk Assessment. Available online at http://www.epa.gov/opp00001/science/comptox-glossary.html (accessed August 29, 2014).
transgenic	Produced from a genetically manipulated egg or embryo; containing genes from another species. NIH's National Center for Biotechnology Information. Transgenic. Available online at http://www.ncbi.nlm.nih.gov/mesh/?term=transgenic (accessed August 29, 2014).
translation	The process of translating the sequence of a messenger RNA (mRNA) molecule to a sequence of amino acids during protein synthesis. The genetic code describes the relationship between the sequence of base pairs in a gene and the corresponding amino acid sequence that it encodes. In the cell cytoplasm, the ribosome reads the sequence of the mRNA in groups of three bases to assemble the protein. NIH's National Human Genome Research Institute. Talking Glossary of Genetic Terms. Available online at http://www.genome.gov/glossary/index.cfm?id=200&textonly=true (accessed August 29, 2014).
translesion synthesis	A mechanism for DNA damage tolerance that allows the DNA replication machinery to move beyond a DNA lesion or abasic site (i.e., a site that lacks a DNA base).
Virtual Tissue (v-Tissues™) models	Computational cross-scale models of cellular organization and emergent functions are used to understand disease progression. Tissues are the clinically relevant level for diagnosing and treating the transition from normal to adverse states in chemical-induced toxicities leading to cancer, immune dysfunction, developmental defects, and more. Currently, *in vivo* rodent experiments are used to evaluate tissue-level effects of altered molecular and cellular function; however, the extrapolation of animal models to humans is often uncertain. v-Tissues™ aims to simulate key molecular and cellular processes computationally in the context of normal tissue biology to (1) help understand complex physiological relationships, and (2) predict adverse effects due to chemicals. As the number of chemicals in consumer products, the workplace, and the environment continues to rise, v-Tissues™ offers a more efficient, effective, and humane approach for evaluating their impact on human health. U.S. Environmental Protection Agency, Computational Toxicology Research Program. http://www.epa.gov/ncct/virtual_tissues/what.html (accessed August 29, 2014).





EPA
United States
Environmental Protection
Agency