Title Page

Manuscript title: Advancing the Next Generation of Risk Assessment

Ila Cote,[1] Melvin Andersen,[2] Gerald Ankley,[1] Stanley Barone, [1] Linda Birnbaum,[3] Kim Boekelheide,[4] Frederic Bois,[5] Lyle Burgoon,[1] Weihsueh A. Chiu,[6] Douglas Crawford-Brown,[7] Kevin M. Crofton,[1] Michael DeVito,[3] Robert Devlin,[1] Stephen Edwards,[1] Kathryn Z. Guyton,[8] Dale Hattis,[9] Richard Judson,[1] Derek Knight,[10] Daniel Krewski,[11] Jason Lambert,[1] Elizabeth Maull,[3] Donna Mendrick,[12] Gregory M. Paoli,[13] Chirag Patel,[14] Edward Perkins,[15] Gerald Poje,[16] Christopher Portier,[17] Ivan Rusyn,[6] Paul Schulte,[18] Anton Simeonov,[19] Martyn Smith,[20] Kristina A. Thayer,[3] Russell S. Thomas,[1] Reuben Thomas, [20] Raymond R. Tice,[3] John Vandenberg,[1] Daniel L. Villeneuve,[1] Scott Wesselkamper,[1] Maurice Whelan,[21] Christine Whittaker,[18] Ronald H. White[22] Menghang Xia,[19] Carole Yauk,[23] Lauren Zeise,[24] Q. Jay Zhao,[1] Robert S. DeWoskin[1]

[1]U.S. Environmental Protection Agency
[2]The Hamner Institutes for Health Sciences
[3]National Institute of Environmental Health Sciences and the National Toxicology Program
[4]Brown University
[5]L'Institut National de l'Environnement Industriel et des Risques
[6]Texas A & M University
[7]University of Cambridge, Department of Land Economy,
[8]International Agency for Cancer Research
[9]Clark University
[10]European Chemicals Agency
[11]University of Ottawa
[12]Food and Drug Administration, National Center for Toxicological Research
[13]Risk Sciences International
[14] Harvard University
[15]Army Corps of Engineers
[16]Grant Consulting Group
[17]Environmental Defense Fund
[18]National Institute for Occupational Safety and Health
[19]National Institutes of Health, National Center for Advancing Translational Science
[20]University of California, Berkeley
[21]European Commission Joint Research Centre, Institute for Health and Consumer Protection
[22]Center for Effective Government
[23]Health Canada
[24]California Environmental Protection Agency, Office of Environmental Health Hazard Assessment

Corresponding author: Ila Cote, U.S. Environmental Protection Agency, Region 8, Room 8152, 1595 Wynkoop St., Denver, Colorado, USA 80202-1129. Telephone: 202-288-9539. Email: cote.ila@epa.gov.

EPA-HQ-2018-000876_0073477

Running title: Advancing the Next Generation of Risk Assessment

Acknowledgments, including grant information

Disclaimer: The views in this article are solely those of the authors, and do not necessarily reflect the views or policies of their agencies or institutes including (but not limited to) the U.S. Environmental Protection Agency (EPA), U.S. Food and Drug Administration (FDA), National Institute for Occupational Safety and Health (NIOSH), California EPA Office of Environmental Health Hazard Assessment (OEHHA), or European Chemicals Agency.

The authors declare they have no actual or potential competing financial interests.

EPA-HQ-2018-000876_0073478

Abstract

Background and Objectives

The Next Generation (NexGen) of Risk Assessment effort is a multiyear collaboration among several agencies and institutions evaluating new, potentially more efficient approaches to environmental health risk assessment. This paper reviews key NexGen findings and identifies strategic research directions.

Methods

The focus was to evaluate how new knowledge from recent advances in molecular, computational, and systems biology might support risk management decisions. Risk assessment prototypes demonstrated application of new data and methods to decision contexts with increasing regulatory impacts. Data types included transcriptomics, genomics, and proteomics; methods included molecular epidemiology and clinical studies, bioinformatic knowledge mining, short-duration bioassays, and quantitative structure activity relationship modeling.

Conclusions

NexGen has fostered extensive discussion among risk scientists and managers and improved confidence in interpreting and applying new data streams in risk-based chemical prioritization and screening, and risk assessment. NexGen has advanced our ability to apply new science by more rapidly identifying chemicals and exposures of potential concern via knowledge mining and use of high- and medium-throughput bioassays; helping characterize mechanisms of action that influence conclusions about causality and exposure-response; and providing conceptual models to evaluate factors affecting susceptibility and cumulative risk.

EPA-HQ-2018-000876_0073479

EPA-HQ-2018-000876_0073480

Advancing the Next Generation of Risk Assessment

Introduction

Background

Advances in molecular and cell biology provide new insights into the etiology of human disease, largely by evaluating molecular events that influence cell function and interactions (Audouze et al. 2013; Hood and Tian 2012; McCullough et al. 2014; McHale et al. 2012; Thomas et al. 2014). High-throughput/high-content (HT/HC) assays and robotic implementation are generating data streams at unprecedented speeds. Computational tools, automated analytical methods (bioinformatics), and systems biology approaches are being developed to organize and interpret the information (Attene-Ramos et al. 2013; Freitas et al. 2014; Hsu et al. 2014; Huang et al. 2014; Judson et al. 2012; Judson et al. 2013; Judson et al. 2014). Toxicity testing and risk assessment will benefit greatly from these advances (Krewski et al. 2014; NRC 2007).

The National Library of Medicine, Tox21, and ToxCast are among the efforts compiling, organizing, managing, and storing these data to better understand determinants of population health (Krewski et al. 2014; NRC 2011) and to help answer such questions as: Which chemicals are environmentally better choices in commerce? Why do individuals and specific subpopulations respond differently to chemical exposures? What happens when people are exposed to low levels of multiple chemicals? How do factors like poverty and preexisting illness influence public health risk? How might evaluating and applying these data, methods, and models support environmental health decisions?

To evaluate how new data types and approaches can inform environmental health

EPA-HQ-2018-000876_0073481

risk assessments, the U.S. Environmental Protection Agency (EPA) collaborated with several U.S. and international agencies and organizations (Supplemental Material, Table S1) to consider the state of science and to develop case studies (illustrative prototypes) demonstrating various approaches that investigators could apply to different risk management problems. Our goal was to provide examples that would promote discussion in the risk assessment, risk management, and stakeholder communities, and that would facilitate the transition from strategy to practical application. This paper summarizes these efforts. A more detailed report is also available (EPA 2014b).

Objectives

Our specific objectives were to test whether these new data sources and methods would help identify specific patterns of molecular events that are (1) associated with impacts of chemical exposures; (2) exposure-dose dependent within the range of environmental exposures; (3) related to such risk factors as genomic variants, chemical and nonchemical stressor coexposures; or (4) useful as improved indicators of adverse health effects and chemical potency. We also considered how new types of assessments might address differing risk management needs or risk context and help develop decision rules for integrating and applying the available data.

Methods

We evaluated and integrated diverse types of data and methods to determine if, and how, advanced biological data would better inform risk assessments.

Preparation for Prototype Development

To establish the foundation for this effort, we (1) worked with EPA risk managers

EPA-HQ-2018-000876_0073482

to identify research needs and develop a strategy for the overall approach (Cote et al. 2012); (2) consulted with experts on the concepts for the prototypes (EPA 2010); (3) held a stakeholder conference to inform the public about upcoming activities and to solicit advice (EPA 2011); and (4) developed a framework articulating the guiding principles for NexGen (Krewski et al. 2014).

Risk Assessments Targeted to Various Decision Contexts

We developed seven prototypes illustrating three decision contexts generally representing environmental challenges risk managers face: (1) major scope decisions, usually regulatory decision-making, generally aimed at nationwide exposures and associated risks; (2) limited scope decisions, often nonregulatory decision-making, generally aimed at limited exposure, hazard, or data situations; and (3) chemical screening and prioritization for further testing, research, or assessment, or for emergency response (Figure 1). These generalized decision contexts do not, and are not meant to, capture all decisions or situational nuances risk managers face.

Study Selection

Establishing systematic review criteria for study selection helps ensure reproducibility, transparency, and scientific acceptability of regulatory actions (DHHS 2014; Meek et al. 2014; NRC 2014; Rhomberg et al. 2013; Rooney et al. 2014).Our criteria were similar to those used for traditional data, augmented with additional criteria specifically applicable to new methodologies (Bourdon-Lacombe et al. In press; McConnell et al. 2014). Rapidly evolving best practices for advanced biology and certain reporting requirements led many initially considered studies to be deemed inadequate for

EPA-HQ-2018-000876_0073483

risk assessment purposes (EPA 2013a).

The Prototypes

Table 1 (Krewski et al. 2014) summarizes the methods considered for the prototypes. Details on the methods and results are provided in EPA (2014a).

Major-scope Assessment Prototypes

Three major-scope prototypes explored how toxicogenomic studies of exposed human populations can inform risk assessment by characterizing early key events in the biological cascade that results in adverse outcomes, biomarkers of exposure and effects, factors contributing to population variability and susceptibility, and the low exposure-response relationship. We developed these prototypes as a proof of concept, and as examples of how new data types could better inform chemical assessments based on robust traditional data.

We evaluated transcriptomic data (epidemiological or clinical) in the range of environmental exposures for three chemicals: (1) benzene and other leukemogens (McHale et al. 2011; McHale et al. 2012; Smith et al. 2011; R Thomas et al. 2012; R Thomas et al. 2013; Thomas et al. 2014); (2) ozone (EPA 2013b; Hatch et al. 2014; McCullough et al. 2014); and (3) polycyclic aromatic hydrocarbons (PAHs), including tobacco smoke and benzo[a]pyrene (DHHS 2014; EPA 2013a; IARC 2010). We also considered genomic, proteomic, and epigenomic data as available, and molecular animal and *in vitro* data for benzene and B[a]P (EPA 2013a; French et al. 2015). We evaluated exposures for benzene of <0.1 to 10 parts per million (ppm) and ozone of 0.5 ppm for 2 hours. We used individual measures of exposure-dose

EPA-HQ-2018-000876_0073484

for benzene and ozone (benzene urinary metabolites and $^{18}O_2$) (Hatch et al. 2014; Vermeulen et al. 2004). For PAH exposures, we used self-reported smoking. The PAH/tobacco smoke prototype focused on pathway mining of existing human microarray data from the Gene Expression Omnibus and ArrayExpress (EMBL-EBI 2015; NCBI 2015). The toxicogenomics data were anchored qualitatively and quantitatively to known health outcomes associated with these chemicals, specifically hematotoxicity and leukemia (benzene and other known leukemogens), lung inflammation and injury (ozone), and lung cancer (PAHs). These data-rich associations therefore enabled us to draw on a wealth of chemical- and disease-specific data to help characterize relationships among upstream molecular changes, downstream cellular events, and public health outcomes. Thus, the potential role of toxicogenomics in hazard identification and dose-response assessment was explored.

Limited-scope Assessment Prototypes

These prototypes explored approaches falling between molecular human clinical and epidemiology studies (above) and *in vitro*, HT screening bioassays (below) in terms of confidence in the data to characterize public health risks, resources expended to collect data, and the number of chemicals that can be evaluated in a given period. We considered three approaches to limited-scope assessment: (1) knowledge mining of large health databases (focusing on human tissue biomonitoring and diabetes data from NHANES [National Health and Nutrition Examination Survey] data) (Bell and Edwards 2015; DeWoskin et al. 2014; EPA 2014b; Patel et al. 2012; Patel et al. 2013a; Thayer et al. 2012); (2) short-duration, *in vivo* exposures using alternative (nonmammalian) species (focusing on the thyroid hormone disruptor mechanism, and zebrafish developmental

EPA-HQ-2018-000876_0073485

outcomes for several hundred chemicals) (Padilla et al. 2012; Perkins et al. 2013; Sipes et al. 2011a; Sipes et al. 2011b; Thienpont et al. 2011; Villeneuve et al. 2014); and (3) short-duration, *in vivo* exposure rodent studies that correlated transcriptomic alterations with cancer and noncancer outcomes as determined in traditional bioassays (Thomas et al. 2011; RS Thomas et al. 2012; RS Thomas et al. 2013a; RS Thomas et al. 2013c). Advantages of the limited-scope approaches compared to HT *in vitro* approaches include intact metabolism and intact cell and tissue interactions, and potential to measure adverse health outcomes, including complex outcomes such as altered behavior and development.

Screening and Prioritization Prototypes

The two screening and prioritization prototypes are (1) quantitative structure activity relationship (QSAR) models and use of analogous chemicals to expand available information (also called "read-across") (EPA 2015b; Golbraikh et al. 2012; OECD 2014; Politi et al. 2014; Wang et al. 2011; Wang et al. 2012a); and (2) *in vitro* cell-based and biochemical-based (including enzymatic and ligand-binding) HT screening assays [focusing on evaluating thyroid hormone disruptors (Cox et al. 2014; Rotroff et al. 2013; Sipes et al. 2011a)]. Of note is that, although QSAR and *in vitro* assays are illustrated separately here, they often are used most effectively in combination. EPA's ToxCast program (Judson et al. 2011; Judson et al. 2012; Judson et al. 2013; Judson et al. 2014) and the multiagency collaborative Tox21 program (Attene-Ramos et al. 2013; Freitas et al. 2014; Hsu et al. 2014; Huang et al. 2014; Tice et al. 2013) provide more information on these approaches. Virtual tissue modeling approaches also are discussed (DeWoskin et al. 2014; Judson et al. 2011; Judson et al. 2012; Judson et al. 2013; Judson et al. 2014;

EPA-HQ-2018-000876_0073486

Kavlock et al. 2012; Kleinstreuer et al. 2014; Knudsen and DeWoskin 2011; Knudsen et al. 2013; Sipes et al. 2013).

Examining Human Variability in Responses

The data to evaluate variability and susceptibility are usually scant. We evaluated several data types to inform this issue: (1) adverse outcome networks to identify mechanistic commonalties among leukemogens and lifestyle factors (diet and stress) that alter leukemia risks (EPA 2014b; IARC 2012; Smith et al. 2011); (2) altered disease incidence in subpopulations having specific genetic polymorphisms (EPA 2014b); (3) data for *in vitro* cells that retain an asthma phenotype in ozone studies (Duncan et al. 2012); (4) correlated measurements of phenotypic differences among diverse subpopulations with different incidences of given exposures [tissue biomonitoring using NHANES (EPA 2014b; Patel et al. 2012; Patel et al. 2013a)]; (5) HT *in vitro* data from cell lines with different genetic backgrounds from the 1000 Genomes effort (Attene-Ramos et al. 2015; Lock et al. 2012; O'Shea et al. 2011); and (6) computational modeling in which variability in parameter values is simulated for differences among subpopulations (Knudsen and DeWoskin 2011; Shah and Wambaugh 2010). See Zeise et al. (2013) for further details.

Results and Discussion

The NexGen prototypes show progress in our understanding of health and disease and help realize the National Research Council's vision embodied in *Toxicity Testing in the 21st Century* (Krewski et al. 2011; NRC 2007). Since this report was published, toxicity testing and risk assessment have begun shifting from the traditional, almost exclusive, use of animal data to

EPA-HQ-2018-000876_0073487

using the new approaches the prototypes demonstrate. The new approaches consider a broader data array, foster mechanistic understanding of adverse effects, and move toward replacing uncertainty factors and extrapolations with data-derived probability distributions.

In each decision context category, new methods and data types were identified that could help inform assessment efforts. Methods illustrated in the screening and prioritization (Tier 1) and limited-scope (Tier 2) prototypes originally were designed for qualitative evaluation of chemicals. New and integrated approaches, however, are being developed to estimate relative potencies and more rapid quantitative toxicity values for use in certain decision contexts.

We used adverse outcome pathways (AOPs) extensively to organize and interpret data for most of the prototypes and regard them as critical for linking molecular events to apical outcomes. The concept of AOPs and networks has gained considerable traction since it was first introduced (Ankley et al. 2010; Davis et al. 2015; Garcia-Reyero 2015; Geer et al. 2010; NAS 2012; Tollefsen et al. 2014; Vinken 2013). We use the terms AOP and AOP network throughput this paper as they are commonly used among many U.S. and European Agencies.

Data quality and reporting are always critically important. Our data searches identified many published studies that we could not use because the data or the reporting was not sufficient for use in health risk assessment. This situation derives from the lag before best practices are fully implemented in the research community and inconsistent application of criteria for data quality and reporting (EPA 2014b; McConnell et al. 2014).

EPA-HQ-2018-000876_0073488

Integrating the available data into a coherent analysis is also a challenge. Supplemental Material, Table S2 presents an "Illustrative Framework for Evidence Integration for New Data Types," focusing on evaluating and integrating evidence and drawing conclusions based on inferences. The table bases evidence integration on previous works (DHHS 2014; EPA 2013a; Meek et al. 2014; NRC 2014; Rooney et al. 2014).

Major-scope Assessment Prototypes (Tier 3)

We designed the Tier 3 prototypes to determine whether new data types could provide results comparable to robust traditional data. We also evaluated whether new data types could add to information robust traditional data sets provide. Support for this hypothesis and several sources of variability are given below (EPA 2013b, 2014b; Esposito et al. 2014; Hatch et al. 2014; McCullough et al. 2014; McHale et al. 2011; McHale et al. 2012; Smith et al. 2011; Thomas et al. 2014). Highlights from the prototypes include:

- AOP networks are useful in predicting specific hazards for benzene and other known leukemogens (hematotoxicity), ozone (lung inflammation and injury), and PAHs (lung cancer). Related chemical and nonchemical stressors (known to cause or exacerbate the same adverse health outcome) were shown to perturb various pathways within the same disease associated network, but do not always affect the same expressed genes or pathway. Hence, overly simplistic descriptions of AOPs could miss the potential for network-level interactions. Evidence for a causal relationship between a specific AOP and adverse effects includes pharmacologic intervention to block identified pathway

EPA-HQ-2018-000876_0073489

changes, use of knock in–out models, or identification of pathway polymorphisms and concomitant amelioration of severity or incidence of the specified adverse outcomes.

- Less well-studied chemicals inducing the same AOP or AOP network could be of concern for concomitant health outcomes. Conversely, lack of an apparent mechanistic link to an adverse outcome might justify downgrading questionable *in vivo* data. Thus, network-level knowledge often is highly valuable to understand causal mechanisms, help integrate evidence, assess potential hazards of well-studied chemicals, provide a basis for cumulative assessment by grouping chemical and nonchemical stressors according to their common AOP network, and evaluate mechanisms underlying human susceptibility (e.g., genetic differences).
- Biomarkers appropriately anchored to AOPs can help elucidate exposure-dose-response relationships as the benzene and ozone prototypes show. Understanding the quantitative relationship of any biomarker to exposure and effect requires substantial study. A most promising application of biomarkers, however, is the ability to measure events of interest directly in environmentally exposed humans—an application revolutionizing epidemiology.
- For benzene, ozone, and theoretically for PAHs, we demonstrated that multiple AOPs developed and progressed with increasing exposures. With benzene, gene and pathway alterations indicative of impaired immune function occur at all exposure levels evaluated (from <0.1 ppm to 10 ppm). At higher concentrations, molecular pathways and effects characteristic of more severe toxicity (apoptosis and cell death) begin to emerge. Data collection over a range of concentrations thus remains essential when evaluating new

EPA-HQ-2018-000876_0073490

data types. Additionally, limited time-course post-exposure data were available for ozone; various adverse outcomes involved in lung injury progressed after exposure, demonstrating the potential dynamic nature of underlying mechanisms (EPA 2013a; McCullough et al. 2014).

- Chemical exposures resulting in adverse outcomes appear to share AOP networks with pathologies of unknown origins (idiopathic or potentially naturally occurring). This suggests that chemically induced events might add to naturally occurring backgrounds of disease via shared mechanisms (EPA 2014b). As NRC (2009) and Crump et al. (1976) discuss, this finding has implications for an assumption of low-dose linearity for cancer and noncancer outcomes at the population level.
- The prototypes helped characterize experimental and organismic factors influencing data interpretation, including experimental variability resulting from differing exposure concentrations, dosimetry, time courses, experimental techniques, experimental paradigms, cell and tissue types, individual genomic profiles, coexposures, and lifestages. Identifying causal events without tight control of variability can be difficult even knowing the adverse outcome, reinforcing the importance for careful experimentation and interpretation when potential outcomes are unknown (EPA 2014b).

Limited-scope Assessment Prototypes (Tier 2)

We designed the Tier 2 prototypes to evaluate data from knowledge mining, alternative species bioassays, and short-term *in vivo* studies for identifying potential hazards, refining mechanistic understanding, and characterizing the relative potencies of thousands of chemicals more rapidly than possible with traditional methods. Confidence

in these data generally ranks between Tier 3 and Tier 1 approaches. Highlights from the prototypes include:

- These approaches are faster and less expensive than the molecular human studies noted above and traditional chronic animal bioassays. Furthermore, unlike the QSAR models and HT screening data (discussed below), the data from *in vivo* studies are from intact organisms with metabolic function, normal architecture (for various cell and tissue types), and normal cell-cell, tissue-tissue interactions. The data also can be used to study more complex system-level adverse outcomes, such as developmental and neurobehavioral outcomes.

- In the data-mining exercises, specific chemical exposures were associated with altered risks for diabetes or prediabetes (e.g., chlorinated organics, heavy metals, selected nutrients). We mined exposure data from NHANES human tissue biomonitoring; NHANES clinically defined incidence. Additional risk factors—multiple chemical exposures and genetic and lifestyle susceptibility traits—were identified (Bell and Edwards 2015; EPA 2014b; Patel et al. 2012; Patel et al. 2013a; Patel et al. 2013b). In one example, 59 percent of people with high levels of cadmium, lead, and arsenic also had markers for diabetes. The data mining results are generally most suitable for hypothesis generation because the output only identifies associations among events in very large data sets. The availability of biomonitoring data and clinical diagnoses in the same individuals, or understanding of mechanisms, however, increases the weight of evidence for these data. Others also have provided traditional and computational data that support a link between chemical exposure and diabetes (Audouze et al. 2013; Dimas et

al. 2014; Inadera 2013; Thayer et al. 2012).

- Two Tier 2 prototypes demonstrated use of short-duration exposure bioassays in alternative species and mammalian species. We evaluated the results with traditional, molecular, and computational approaches. Collectively, these bioassays successfully identified exposure concentrations associated with adverse outcomes and related key events and AOP network alterations linked to adverse effects. These prototypes provided data on complex mechanistic behaviors, effects of mixtures, and species-to-species similarities and differences, illustrating how these data could be used to evaluate potential hazards and chemical potencies (Ankley and Gray 2013; Padilla et al. 2012; Painter et al. 2014; Perkins et al. 2013; RS Thomas et al. 2013b; RS Thomas et al. 2013c).

Screening and Prioritization Prototypes (Tier 1)

For the first time, new approaches are being used that can evaluate vast numbers of chemicals relatively rapidly. For example, tens of thousands of chemicals the European Regulation on Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH) Legislation covers are being evaluated using QSAR and new types of bioassays. The U.S. Tox21 program is screening approximately 8,500 chemicals using innovative robotic technology and *in vitro* bioassays (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." The QSAR models (Goldsmith et al. 2012; Venkatapathy and Wang 2013; 2012a; Wang et al. 2012b) and HT *in vitro* bioassays were used to illustrate the rapid

EPA-HQ-2018-000876_0073493

successful screening and prioritization of chemicals (Judson et al. 2013; Kavlock et al. 2012; Rusyn et al. 2012; Sipes et al. 2013; Tice et al. 2013). Additional insights include:

- An essential element to evaluating and applying HT data within the risk paradigm is dose characterization. Researchers are developing methods using reverse dosimetry to extrapolate bioactive concentrations in *in vitro* test systems to the comparable doses for *in vivo* exposure to rodents (or other test species) or to humans (*in vitro*-to-*in vivo* extrapolation [IVIVE]) (Hubal 2009; Rotroff et al. 2010; Wetmore et al. 2012; Wetmore et al. 2013). IVIVE extrapolation supports quantitative comparisons of *in vitro* toxicity results with *in vivo* bioassay results for estimating dose-response in human exposures.

- QSAR, *in vitro,* and *in silico* methods, are proving useful for screening and ranking large numbers of chemicals for further assessment and urgent-response situations where traditional data are lacking. Current estimates of human disease risks based exclusively on QSAR and *in vitro* HT screening generally are too uncertain; *in silico* models, however, are improving our understanding of these data. Insights into underlying mechanisms of toxicity, and the factors that might contribute to population variability in response to chemical exposure (Lock et al. 2012; O'Shea et al. 2011), are also progressing from these data streams and increasing their utility for understanding risks.

Caveats Pertaining to Applying New Data Types in Risk Assessment

Exposure and adverse outcomes often can be associated with hundreds to thousands of gene changes, not all of which are causal (Mendrick 2011). Associative data can "suggest" 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 causative relationships between key events in AOP networks and consequent adverse health outcomes; or combinations of consistent, coherent traditional and new data types. The prototypes demonstrated how different types of evidence in each decision support category might be characterized with respect to causality and evidence integration (EPA 2013b, 2014a; NRC 2014). Additional caveats are described below. Many of these concerns apply to traditional, as well as new data types.

- Cell type, tissue, individual, subpopulation, strain, 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 likely is due, at least in part, to dosimetry, epigenomic differences, and genomic plasticity, which assessments should consider whenever feasible.
- For many chemicals, metabolism is critical to toxicity. That most HT *in vitro* test systems have limited or no metabolic competence should be considered. Although researchers are evaluating various approaches to add or enhance metabolic capability, satisfactory solutions for routine screening of larger numbers of chemicals are not yet available. Consequently, although positive results are informative, negative results should not yet be interpreted as a lack of toxicity.
- Molecular profiles appear to be both dose and time dependent. Predicting adverse outcomes based only on "snapshots" of biological events can therefore be challenging. Focusing on profiles associated with environmentally relevant exposures should improve

predictions. Some signatures do appear stable over time, however, and might also serve as reliable indicators of chronic outcomes (RS Thomas et al. 2013c).

- Gene expression data contain much uncertainty, as messenger ribonucleic acid expression levels cannot be used to infer protein activity directly. These data alone could be suitable for ranking and screening and used in assessments to complement other mechanistic data.

- Our current ability to monitor multiple molecular processes (genomics, transcriptomics, proteomics, and epigenomics) in a single study is very limited, primarily due to cost. This hampers biological integration and limits our understanding of how chemicals influence complex biological systems.

- Only a few chemicals represented in the current literature have biological data adequate to support regulatory risk assessments, due primarily to experimental design and reporting issues. This limitation reinforces the need for systematic review.

- A major challenge in using molecular data in risk assessment is how to use the data to improve predictions of adverse effects in humans. For example, how do changes in molecular events affect cells, changes in cells affect tissues and organs, and changes in organs affect the whole body? Researchers are collecting large amounts of HT/HC screening data on molecular-level effects, and the body of information on diseases and disease outcomes is substantial. Very sparse chemical-specific data are available, however, on intermediate levels of organization and on the sequence of cellular-level disruption of normal biology to effects at higher organizational levels. Even so, tremendous strides are being made in generating disease-specific information.

EPA-HQ-2018-000876_0073496

- Characterizing population response variability among individuals is a major challenge, given the many sources of inherent biological variability (e.g., genetic differences) and extrinsic influences (e.g., lifestyle, poverty, multiple chemical exposures). Each chemical exposure-health outcome pair involves combinations of these sources, and different decision contexts present distinct needs regarding the identification—and extent of characterization—of interindividual variability in the human population. New approaches to examining variability in responses include (1) computational modeling, in which variability in parameter values is simulated and differences among subpopulations are explored (Diaz Ochoa et al. 2012; Knudsen and DeWoskin 2011; Shah and Wambaugh 2010); (2) HT *in vitro* data analysis of cell lines with different genetic backgrounds from the 1000 Genomes effort (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 (French et al. 2015; NIEHS 2015c); (4) comprehensive scanning of gene coding regions in diverse individuals to examine the relationships among environmental exposures, interindividual sequence variation in human genes, and population disease risks (Mortensen and Euling 2013; NIEHS 2015a); (5) genome-wide association studies to uncover genomic loci that might contribute to risk of disease (NHGRI 2015; Wright et al. 2012); and (6) association studies correlating phenotypic differences among diverse populations with expression patterns for groups of genes based on coexpression (Friend 2013; Patel et al. 2012; Patel et al. 2013a; Weiss et al. 2012). Additionally, understanding of the contribution of epigenomics to disease is advancing rapidly (Ghantous et al. 2015).

EPA-HQ-2018-000876_0073497

- Verifying toxicity testing schemes and computational models that are more efficient is essential for using these new data and approaches for risk-based decisions. Central to this effort are a framework and criteria for determining whether the new data types are adequate for various types of decisions. The level of certainty needed in the data varies with their intended use because inaccurate results have increasing consequences and costs as decisions progress from screening, to further testing, to what safe chemical levels are, to what regulatory (or mitigation) actions should be taken (Crawford-Brown 2013). Traditional "validation" approaches that evaluate conventional assay and testing structures do not adequately address the potential uses of these new data and methods and would require years to implement (Judson et al. 2013). Thus, as the technology for rapid, efficient, robust hazard testing advances, the verification process also must advance to ensure confidence in their use. Clear and transparent articulation of these decision considerations are essential to the acceptance of, and support for, assessment results and in the overall evidence integration.

Based on the lessons learned in the NexGen program and elsewhere, several new types of high- and medium-throughput assessments are being advanced. Table 2 shows how characteristics of "fit-for-purpose" assessments could be tailored to support three illustrative decision-context categories. The table lists potential uses for NexGen assessments, data sources and types in different assessment categories, exposure paradigms used, incorporation of toxicokinetics, use of traditional data, hazard characterization, potency metrics, inferences drawn about the causal associations between exposures and adverse outcomes, the numbers of chemicals that can be assessed, and the time to conduct any given assessment.

Research Needs

Enhancing our understanding of complex chemical and biological interactions at various levels of biological organization requires integrating computational research with strategic laboratory studies to advance available models and accelerate application of new data in risk assessment. We suggest focusing on the following specific areas:

- Development of reliable, molecular biomarkers and bioindicators representing a wide variety of chemical exposures and key events of pathogenesis for building confidence in the characterization of key events used to construct an AOP.
- Identification and understanding of AOP network interactions among different levels of organization for observed key events (genes, proteins, cells, tissues, organs, individuals, populations and communities), including characterization of compensatory responses and their prognostic value for different adverse outcomes or disease states.
- Collection of data and development of methods for reverse toxicokinetics to extrapolate concentrations used in cellular and cell-free systems to *in vivo* tissue doses and exposures; nonaqueous *in vitro* exposure methods for chemicals present as gases or as airborne particles; and adjusting for intra- and interspecies differences when assessing potential human effects based on nonhuman toxicity data.
- Approaches for grouping chemical and nonchemical stressors based on common key events within AOPs to enable cumulative risk assessment; considerations for source apportionment with respect to exposures for cumulative risk assessment.
- Evaluation of individual human variability due to lifestage vulnerabilities, genetic

EPA-HQ-2018-000876_0073499

differences, preexisting disease and exposure, or adaptive and compensatory capabilities; and development of techniques to incorporate this variability into population-level risk assessment.

Conclusions

A revolution in molecular, computational, and systems biology is rapidly advancing our understanding of what causes disease and who becomes affected, and the role of environmental factors on public health. This information is just beginning to result in innovative, more efficient approaches to toxicity testing and risk assessment. This paper summarizes recent, multiorganizational efforts to understand and apply emerging science in a transparent and scientifically robust manner. We anticipate these novel methods will provide a more complete understanding of the biological underpinnings of health risks and, also, methods and data to help evaluate the tens of thousands of unaddressed chemicals in the nation (EPA 2015a). The overarching challenge to risk assessors is to obtain and interpret sufficient data for quick and efficient assessment to support decisions that protect public health and the environment. The ultimate goal is to develop safer chemicals and to better manage risks to public health and the environment. The prototypes demonstrate how new data can be used to help address these challenges.

Ongoing efforts to advance toxicity testing and risk assessment include:

- Thousands of chemicals, previously having no or very limited traditional data, are being assessed based on similarities in physical-chemical structure to known toxicants (QSAR modeling) and on the results of rapid, robotically conducted *in vitro* bioassays. These evaluations will help prioritize testing, research, and assessment, and in emergency

EPA-HQ-2018-000876_0073500

response situations.

- Hundreds of chemicals are being evaluated by using computational analyses of large primary databases held in public repositories and by identifying the most important findings in the burgeoning literature. These efforts are playing a central role in developing knowledge about the potential toxicity of chemicals and the causes of disease. These approaches, in combination with high throughput approaches, could be used to support limited scope assessments or to augment robust traditional data-based assessments.
- Developing, innovative, targeted testing approaches that combine short-duration in vivo bioassays and HT approaches will provide even more robust information for testing and assessment.
- Finally, a variety of new methods are addressing the formidable challenges of characterizing cumulative effects from exposure to multiple chemical and nonchemical stressors, susceptible subpopulations, and low-dose responses, primarily based on improving mechanistic understanding of adverse health effects.

Near-term efforts include developing additional prototypes for public input and peer review and providing more opportunities to solicit stakeholder comments and participation. EPA is developing a verification process for new methods and data types that focus on integrating the evidence into various decision contexts for use by risk assessors. The goal is to increase confidence for using these new approaches in risk assessment. Significant scientific gaps identified in the completed and ongoing prototypes are helping guide future research plans.

EPA-HQ-2018-000876_0073501

We anticipate the prototype demonstrations will help overcome the significant logistical and methodological challenges in interpreting and using these new data and methods in risk assessment. For now, major chemical assessments will continue to be driven primarily by traditional data but with increasing augmentation with the new types of data. EPA risk managers and the risk assessment community at large will continue to be informed of the new tools and methods being developed with an emphasis on high-quality science and transparency. Historically difficult risk assessment questions that this new and emerging knowledge are likely to inform include: Why do individual and specific populations respond differently to environmental exposures? Why are children at greater risk for certain exposures and effects? What happens when people are exposed to mixtures that contain very low levels of individual chemicals, such as those commonly found in the environment? How do other environmental factors like poverty and preexisting health conditions alter the response to chemical exposures? These are just some of issues that NexGen assessments will help address to improve the identification of safer chemicals and reduce risk from exposures to hazardous chemicals in the environment.

EPA-HQ-2018-000876_0073502

References

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. Environ Toxicol Chem 29:730–741.

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. Environ Toxicol Chem 32:1084–1087.

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 18:716–273.

Attene-Ramos MS, Huang R, Michael S, Witt KL, Richard A, Tice RR, et al. 2015. Profiling of the Tox21 chemical collection for mitochondrial function to identify compounds that acutely decrease mitochondrial membrane potential. Environ Health Perspect 123:49–56.

Audouze K, Brunak S, Grandjean P. 2013. A computational approach to chemical etiologies of diabetes. Sci Rep 3:2712; doi:10.1038/srep02712.

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. Regul Toxicol Pharmacol 68:212–221.

Bell SM, Edwards SW. 2015. Identification and prioritization of relationships between environmental stressors and adverse human health impacts. Environ Health Perspect; doi: 10.1289/ehp.1409138 [Online 10 April 2015].

Bourdon-Lacombe JA, Deveau M, Husain M, Moffat ID, Auerbach S, Krewski D, et al. In press. Technical guide for applications of gene expression profiling in human health risk assessment of environmental chemicals. Regul Toxicol Pharmacol.

Cooper WA, Lam DC, O'Toole SA, Minna JD. 2013. Molecular biology of lung cancer. J Thorac Dis 5 Suppl 5:S479–490.

Cote I, Anastas PT, Birnbaum LS, Clark RM, Dix DJ, Edwards SW, et al. 2012. Advancing the next generation of health risk assessment. Environ Health Perspect 120:1499–1502.

Cox LA, Popken D, Marty MS, Rowlands JC, Patlewicz G, Goyak KO, et al. 2014. Developing scientific confidence in HTS-derived prediction models: lessons learned from an endocrine case study. Regul Toxicol Pharmacol 69:443–450.

Crawford-Brown D. 2013. The role of advanced biological data in the rationality of risk-based regulatory decisions. J Environ Prot (Irvine, Calif) 4:238–249.

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.

Davis AP, Grondin CJ, Lennon-Hopkins K, Saraceni-Richards C, Sciaky D, King BL, et al. 2015. The Comparative Toxicogenomics Database's 10th year anniversary: update 2015. Nucleic Acids Res 43:D914–920.

DeWoskin RS, Knudsen TB, Shah I. 2014. Virtual models (vM). In: Encyclopedia of Toxicology, 3rd Edition, Vol. 4 (Wexler P, ed). London, UK:Elsevier, Inc., 948–957.

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. Available: http://www.surgeongeneral.gov/library/reports/50-years-of-progress/exec-summary.pdf [accessed 28 April 2015].

EPA-HQ-2018-000876_0073503

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; doi:10.3389/fphar.2012.00204.
Dimas AS, Lagou V, Barker A, Knowles JW, Magi R, Hivert MF, et al. 2014. Impact of type 2 diabetes susceptibility variants on quantitative glycemic traits reveals mechanistic heterogeneity. Diabetes 63:2158–2171.
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. Am J Respir Crit Care Med 185:A4291; doi:10.1164/ajrccm-conference.2012.185.1_MeetingAbstracts.A4291.
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.
EMBL-EBI (European Molecular Biology Laboratory-European Bioinformatics Institute). 2015. ArrayExpress. Available: http://www.ebi.ac.uk/ [accessed 28 April 2015].
EPA (U.S. Environmental Protection Agency). 2010. Advancing the Next Generation (NexGen) of Risk Assessment: The Prototypes Workshop. EPA/600/R-11/100. Research Triangle Park, NC:U.S. EPA. Available: http://www.epa.gov/risk/nexgen/docs/NexGen-Prototypes-Workshop-Summary.pdf [accessed 28 April 2015].
EPA (U.S. Environmental Protection Agency). 2011. Advancing the Next Generation (NexGen) of Risk Assessment: Public Dialogue Conference Report. EPA/600/R-11/101. Washington, DC:U.S. EPA. Available: http://www.epa.gov/risk/nexgen/docs/NexGen-Public-Conf-Summary.pdf [accessed 28 April 2015].
EPA (U.S. Environmental Protection Agency). 2013a. IRIS Toxicological Review of Benzo[a]pyrene (Public Comment Draft). Washington, DC:EPA. Available: http://cfpub.epa.gov/ncea/iris_drafts/recordisplay.cfm?deid=66193 [accessed 28 April 2015].
EPA (U.S. Environmental Protection Agency). 2013b. Integrated Science Assessment of Ozone and Related Photochemical Oxidants (Final Report). EPA/600/R-10/076F. Washington, DC:U.S. EPA. Available: http://cfpub.epa.gov/ncea/isa/recordisplay.cfm?deid=247492 [accessed 28 April 2015].
EPA (U.S. Environmental Protection Agency). 2014a. Advancing the Next Generation (NexGen) of Risk Assessment: Conferences and Workshops. Washington, DC. Available: http://www.epa.gov/risk/nexgen/workshops.htm [accessed August 29, 2014].
EPA 2014b. Next Generation Risk Assessment: Incorporation of Recent Advances in Molecular, Computational, and Systems Biology (Final Report). Washington, DC:U.S. EPA. Available: http://cfpub.epa.gov/ncea/risk/recordisplay.cfm?deid=286690 [accessed 28 April 2015].
EPA (U.S. Environmental Protection Agency). 2015a. Basic Information about the Toxic Substance Control Act Inventory. Available: http://www.epa.gov/oppt/existingchemicals/pubs/tscainventory/basic.html#background [accessed 28 April 2015].
EPA (U.S. Environmental Protection Agency). 2015b. Quantitative Structure Activity Relationship. Available: http://www.epa.gov/nrmrl/std/qsar/qsar.html [accessed 28 April 2015].
Esposito S, Tenconi R, Lelii M, Preti V, Nazzari E, Consolo S, et al. 2014. Possible molecular mechanisms linking air pollution and asthma in children. BMC Pulm Med 14:31; doi: 10.1186/1471-2466-14-31.

Freitas J, Miller N, Mengeling BJ, Xia M, Huang R, Houck K, et al. 2014. Identification of thyroid hormone receptor active compounds using a quantitative high-throughput screening platform. Curr Chem Genomics Transl Med 8:36–46.
French JE, Gatti DM, Morgan DL, Kissling GE, Shockley KR, Knudsen GA, et al. 2015. Diversity outbred mice identify population-based exposure thresholds and genetic factors that influence benzene-induced genotoxicity. Environ Health Perspect 123:237–245.
Friend S. Scientific opportunities from heterogeneous biological data analysis: overcoming complexity. Presentation. National Academy of Sciences Meeting: Integrating Environmental Health Data to Advance Discovery, 10–11 January 2013. Washington, DC.
Garcia-Reyero N. 2015. Are adverse outcome pathways here to stay? Environ Sci Technol 49:3–9.
Geer LY, Marchler-Bauer A, Geer RC, Han L, He J, He S, et al. 2010. The NCBI BioSystems database. Nucleic Acids Res 38:D492–496.
Ghantous A, Hernandez-Vargas H, Byrnes G, Dwyer T, Herceg Z. 2015. Characterising the epigenome as a key component of the fetal exposome in evaluating in utero exposures and childhood cancer risk. Mutagenesis; doi: 10.1289/ehp.1409138 [Online 26 February 2015].
Golbraikh A, Wang XS, Zhu H, Tropsha A. 2012. Predictive QSAR modeling: methods and applications in drug discovery and chemical risk assessment. In: Handbook of Computational Chemistry, (Leszczynski J, ed). New York, NY:Springer, 1309–1342.
Goldsmith MR, Peterson SD, Chang DT, Transue TR, Tornero-Velez R, Tan YM, et al. 2012. Informing mechanistic toxicology with computational molecular models. Methods Mol Biol 929:139–165.
Hatch GE, Duncan KE, Diaz-Sanchez D, Schmitt MT, Ghio AJ, Carraway MS, et al. 2014. Progress in assessing air pollutant risks from *in vitro* exposures: matching ozone dose and effect in human airway cells. Toxicol Sci 141:198–205.
Hatzimichael E, Crook T. 2013. Cancer epigenetics: New therapies and new challenges. J Drug Deliv 2013:529312; doi: 10.1155/2013/529312:529312.
Hood L, Tian Q. 2012. Systems approaches to biology and disease enable translational systems medicine. Genomics Proteomics Bioinformatics 10:181–185.
Hsu CW, Zhao J, Huang R, Hsieh JH, Hamm J, Chang X, et al. 2014. Quantitative high-throughput profiling of environmental chemicals and drugs that modulate farnesoid X receptor. Sci Rep 4:6437.
Huang R, Sakamuru S, Martin MT, Reif DM, Judson RS, Houck KA, et al. 2014. Profiling of the Tox21 10K compound library for agonists and antagonists of the estrogen receptor alpha signaling pathway. Sci Rep 4:5664.
Hubal EA. 2009. Biologically relevant exposure science for 21st century toxicity testing. Toxicol Sci 111:226–232.
IARC. 2010. Some non-heterocyclic polycyclic aromatic hydrocarbons and some related exposures. IARC Monogr Eval Carcinog Risks Hum 92:1–853.
IARC. 2012. Benzene. IARC Monogr Eval Carcinog Risk Hum 100F:249–294.
Inadera H. 2013. Developmental origins of obesity and type 2 diabetes: molecular aspects and role of chemicals. Environ Health Prev Med 18:185–197.
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. Chem Res Toxicol 24:451–462.
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. Int J Mol Sci 13:1805–1831.
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. ALTEX 30:51–56.
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 Clin Pharmacol Toxicol 115:69–76.
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. Chem Res Toxicol 25:1287–1302.
Kleinstreuer NC, Yang J, Berg EL, Knudsen TB, Richard AM, Martin MT, et al. 2014. Phenotypic screening of the ToxCast chemical library to classify toxic and therapeutic mechanisms. Nat Biotechnol 32:583–591.
Knudsen T, DeWoskin R. 2011. Systems modeling in developmental toxicity. In: Handbook of Systems Toxicology, (Casciano D., Sahu S, eds). Chichester, UK:John Wiley & Sons, Ltd., 599–617.
Knudsen T, Martin M, Chandler K, Kleinstreuer N, Judson R, Sipes N. 2013. Predictive models and computational toxicology. Methods Mol Biol 947:343–374.
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.
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. Toxicol Sci 126:578–588.
McConnell ER, Bell SM, Cote I, Wang RL, Perkins EJ, Garcia-Reyero N, et al. 2014. Systematic Omics Analysis Review (SOAR) tool to support risk assessment. PloS One 9:e110379; doi: 10.1371/journal.pone.0110379.
McCullough SD, Duncan KE, Swanton SM, Dailey LA, Diaz-Sanchez D, Devlin RB. 2014. 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 51:426–435.
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.
McHale CM, Zhang L, Lan Q, Vermeulen R, Li G, Hubbard AE, et al. 2011. Global gene expression profiling of a population exposed to a range of benzene levels. Environ Health Perspect 119:628–634.
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.
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.
Mendrick DL. 2011. Transcriptional profiling to identify biomarkers of disease and drug response. Pharmacogenomics 12:235–249.
Moreno-Macias H, Dockery DW, Schwartz J, Gold DR, Laird NM, Sienra-Monge JJ, et al. 2013.

Ozone exposure, vitamin C intake, and genetic susceptibility of asthmatic children in Mexico City: a cohort study. Respir Res 14:14; doi:10.1186/1465-9921-14-14.
Mortensen HM, Euling SY. 2013. Integrating mechanistic and polymorphism data to characterize human genetic susceptibility for environmental chemical risk assessment in the 21st century. Toxicol Appl Pharmacol 271:395–404.
NAS (National Academy of Sciences). 2012. Emerging Science for Environmental Health Decisions; Workshop − Systems Biology – Informed Risk Assessment. National Academy of Sciences Building. 2101 Constitution Avenue, NW (between 21st and C Streets, NW). Washington, DC. June 14-15. Available: http://nas-sites.org/emergingscience/meetings/omics-informed-risk-assessment/ [accessed
NCBI (National Center for Biotechnology Information). 2015. Gene Expression Omnibus. Available: http://www.ncbi.nlm.nih.gov/geo/ [accessed 28 April 2015].
NHGRI (National Human Genome Research Institute). 2015. A Catalog of Published Genome-Wide Association Studies. Available: http://www.genome.gov/gwastudies/ [accessed 28 April 2015].
NIEHS (National Institute of Environmental Health Sciences). 2015a. NIEHS Single Nucleotide Polymorphisms (SNPs) Environmental Genome Project. Available: http://egp.gs.washington.edu/ [accessed 28 April 2015].
NIEHS (National Institute of Environmental Health Sciences). 2015b. Comparative Toxicogenomic Database (CTD)™. Available: http://ctdbase.org/ [accessed 28 April 2015].
NIEHS (National Institute of Environmental Health Sciences). 2015c. Host Susceptibility Program. Available: http://grants.nih.gov/grants/guide/notice-files/NOT-ES-07-009.html [accessed 28 April 2015].
NRC (National Research Council). 2007. Toxicity Testing in the 21st Century: A Vision and a Strategy. Washington, DC:The National Academies Press. Available: http://dels.nas.edu/resources/static-assets/materials-based-on-reports/reports-in-brief/Toxicity_Testing_final.pdf [accessed 28 April 2015].
NRC (National Research Council). 2009. Science and Decisions: Advancing Risk Assessment. Washington, DC:The National Academies Press. Available: http://www.nap.edu/catalog/12209.html [accessed 28 April 2015].
NRC (National Research Council). 2011. Predicting Later-Life Outcomes of Early-Life Exposures.The National Academies Press. Available: http://nas-sites.org/emergingscience/files/2011/05/inutero_final_April2011.pdf [accessed 28 April 2015].
NRC (National Research Council). 2014. Review of EPA's Integrated Risk Information System (IRIS) Process. Washington, DC:The National Academies Press. Available: http://www.nap.edu/catalog.php?record_id=18764 [accessed 28 April 2015].
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. Toxicol Sci 119:398–407.
OECD (Organisation for Economic Cooperation and Development). 2014. The OECD QSAR Toolbox. Version 3.0. Available: http://www.oecd.org/env/chemicalsafetyandbiosafety/assessmentofchemicals/theoecdqsartoolbox.htm [accessed 29 August 2014].
OECD (Organisation for Economic Cooperation and Development). 2015a. Other Activities on Molecular Screening and Toxicogenomics. Available: http://www.oecd.org/env/ehs/testing/toxicogenomics.htm [accessed 28 April 2015].
OECD (Organisation for Economic Cooperation and Development). 2015b. Grouping of

Chemicals: Chemical Categories and Read-Across. Available: http://www.oecd.org/chemicalsafety/risk-assessment/groupingofchemicalschemicalcategoriesandread-across.htm [accessed 28 April 2015].
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. Reprod Toxicol 33:174–187.
Painter K, McConnell ER, Sahasrabudhe S, Burgoon L, Powers CM. 2014. What do the data show? Knowledge map development for comprehensive environmental assessment. Integr Environ Assess Manag 10:37–47.
Patel CJ, Chen R, Butte AJ. 2012. Data-driven integration of epidemiological and toxicological data to select candidate interacting genes and environmental factors in association with disease. Bioinformatics 28:i121–126.
Patel CJ, Cullen MR. 2012. Genetic variability in molecular responses to chemical exposure. EXS 101:437–457.
Patel CJ, Chen R, Kodama K, Loannidis JP, Butte AJ. 2013a. Systematic identification of interaction effects between genome-and environment-wide associations in type 2 diabetes mellitus. Hum Genet 132:495–508.
Patel CJ, Sivadas A, Tabassum R, Preeprem T, Zhao J, Arafat D, et al. 2013b. Whole genome sequencing in support of wellness and health maintenance. Genome Med 5:58; doi: 10.1186/gm462.
Patlewicz G, Ball N, Booth ED, Hulzebos E, Zvinavashe E, Hennes C. 2013. Use of category approaches, read-across and (Q)SAR: general considerations. Regul Toxicol Pharmacol 67:1–12.
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. Environ Health Perspect 121:1002–1010.
Politi R, Rusyn I, Tropsha A. 2014. Prediction of binding affinity and efficacy of thyroid hormone receptor ligands using QSAR and structure-based modeling methods. Toxicol Appl Pharmacol 280:177–189.
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. Crit Rev Toxicol 43:753–784.
Rooney AA, Boyles AL, Wolfe MS, Bucher JR, Thayer KA. 2014. Systematic review and evidence integration for literature-based environmental health science assessments. Environ Health Perspect 122:711–718.
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. Toxicol Sci 117:348–358.
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.
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. Toxicol Sci 127:1–9.
Schlenk RF, Dohner K, Krauter J, Frohling S, Corbacioglu A, Bullinger L, et al. 2008. Mutations and treatment outcome in cytogenetically normal acute myeloid leukemia. N Engl J Med 358:1909–1918.

Shah I, Wambaugh J. 2010. Virtual tissues in toxicology. J Toxicol Environ Health B Crit Rev 13:314–328.
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. Exp Mol Med 43:374–378.
Sillé FC, Thomas R, Smith MT, Conde L, Skibola CF. 2012. Post-GWAS functional characterization of susceptibility variants for chronic lymphocytic leukemia. PloS One 7:e29632; doi: 10.1371/journal.pone.0029632:e29632.
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. Toxicol Sci 124:109–127.
Sipes NS, Padilla S, Knudsen TB. 2011b. Zebrafish: as an integrative model for twenty-first century toxicity testing. Birth Defects Res C Embryo Today 93:256–267.
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. Chem Res Toxicol 26:878–895.
Smith MT, Zhang L, McHale CM, Skibola CF, Rappaport SM. 2011. Benzene, the exposome and future investigations of leukemia etiology. Chem Biol Interact 192:155–159.
Thayer KA, Heindel JJ, Bucher JR, Gallo MA. 2012. Role of environmental chemicals in diabetes and obesity: a National Toxicology Program workshop review. Environ Health Perspect 120:779–789.
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. Environ Sci Technol 45:7525–7532.
Thomas R, Phuong J, McHale CM, Zhang L. 2012. Using bioinformatic approaches to identify pathways targeted by human leukemogens. Int J Environ Res Public Health 9:2479–2503.
Thomas R, McHale CM, Lan Q, Hubbard AE, Zhang L, Vermeulen R, et al. 2013. Global gene expression response of a population exposed to benzene: a pilot study exploring the use of RNA-sequencing technology. Environ Mol Mutagen 54:566–573.
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; doi: 10.1371/journal.pone.0091828.
Thomas RS, Clewell HJ, 3rd, Allen BC, Wesselkamper SC, Wang NC, Lambert JC, et al. 2011. Application of transcriptional benchmark dose values in quantitative cancer and noncancer risk assessment. Toxicol Sci 120:194–205.
Thomas RS, Clewell HJ, 3rd, Allen BC, Yang L, Healy E, Andersen ME. 2012. Integrating pathway-based transcriptomic data into quantitative chemical risk assessment: a five chemical case study. Mutat Res 746:135–143.
Thomas RS, Himmelstein MW, Clewell HJ, 3rd, Yang Y, Healy E, Black MB, et al. 2013a. Cross-species transcriptomic analysis of mouse and rat lung exposed to chloroprene. Toxicol Sci 131:629–640.
Thomas RS, Philbert MA, Auerbach SS, Wetmore BA, DeVito MJ, Cote I, et al. 2013b. Incorporating new technologies into toxicity testing and risk assessment: moving from 21st century vision to a data-driven framework. Toxicol Sci 136:4–18.
Thomas RS, Wesselkamper SC, Wang NC, Zhao QJ, Petersen DD, Lambert JC, et al. 2013c. Temporal concordance between apical and transcriptional points of departure for chemical risk

assessment. Toxicol Sci 134:180–194.
Tice RR, Austin CP, Kavlock RJ, Bucher JR. 2013. Improving the human hazard characterization of chemicals: a Tox21 update. Environ Health Perspect 121:756–765.
Tollefsen KE, Scholz S, Cronin MT, Edwards SW, de Knecht J, Crofton K, et al. 2014. Applying Adverse Outcome Pathways (AOPs) to support Integrated Approaches to Testing and Assessment (IATA). Regul Toxicol Pharmacol 70:629–640.
Vawda S, Mansour R, Takeda A, Funnell P, Kerry S, Mudway I, et al. 2014. Associations between inflammatory and immune response genes and adverse respiratory outcomes following exposure to outdoor air pollution: a HuGE systematic review. Am J Epidemiol 179:432–442.
Venkatapathy R, Wang NC. 2013. Developmental toxicity prediction. Methods Mol Biol 930:305–340.
Vermeulen R, Li G, Lan Q, Dosemeci M, Rappaport SM, Bohong X, et al. 2004. Detailed exposure assessment for a molecular epidemiology study of benzene in two shoe factories in China. Ann Occup Hyg 48:105–116.
Villeneuve D, Volz DC, Embry MR, Ankley GT, Belanger SE, Leonard M, et al. 2014. Investigating alternatives to the fish early-life stage test: a strategy for discovering and annotating adverse outcome pathways for early fish development. Environ Toxicol Chem 33:158–169.
Vinken M. 2013. The adverse outcome pathway concept: a pragmatic tool in toxicology. Toxicology 312:158–165.
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. Regul Toxicol Pharmacol 59:215–226.
Wang N, Jay Zhao Q, Wesselkamper SC, Lambert JC, Petersen D, Hess-Wilson JK. 2012a. Application of computational toxicological approaches in human health risk assessment. I. A tiered surrogate approach. Regul Toxicol Pharmacol 63:10–19.
Wang N, Rice GE, Teuschler LK, Colman J, Yang RS. 2012b. An *in silico* approach for evaluating a fraction-based, risk assessment method for total petroleum hydrocarbon mixtures. J Toxicol 2012:410143; doi: 10.1155/2012/410143.
Weiss JN, Karma A, MacLellan WR, Deng M, Rau CD, Rees CM, et al. 2012. "Good enough solutions" and the genetics of complex diseases. Circ Res 111:493–504.
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. Toxicol Sci 125:157–174.
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. Toxicol Sci 132:327–346.
Wright FA, Shabalin AA, Rusyn I. 2012. Computational tools for discovery and interpretation of expression quantitative trait loci. Pharmacogenomics 13:343–352.
Zeise L, Bois FY, Chiu WA, Hattis D, Rusyn I, Guyton KZ. 2013. Addressing human variability in next-generation human health risk assessments of environmental chemicals. Environ Health Perspect 121:23–31.
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. PloS One 7:e46974; doi: 10.1371/journal.pone.0046974.

EPA-HQ-2018-000876_0073510

Table 1. Prototype use of new scientific tools and techniques (adapted from Krewski et al. 2014)

Decision-context category	Screening and prioritization	Limited-scope assessments	Major-scope assessments
Hazard identification and dose-response assessment methods			
Quantitative structure activity relationship models	■	■	
Pathway 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		■	■

EPA-HQ-2018-000876_0073511

Table 2. Possible characteristics of fit-for-purpose assessments matched to illustrative decision-context categories.

Characteristics	Illustrative decision-context categories		
	Screening and prioritization	**Limited-scope assessments**	**Major-scope assessments**
Uses of NexGen assessments	Screening chemicals with no data other than QSAR or HT data, e.g., • Queuing for research, testing, or assessment • Urgent or emergency response	Generally nonregulatory decision-making, e.g., • Urban air toxics • Potential water contaminants • Hazardous waste and superfund chemicals • Urgent or emergency response	Often regulatory decision-making, e.g., • National risk assessments • Community risk assessment • Special problems of national concern
Data sources	EPA databases such as ACToR and ToxCast; PubChem	Large public data and literature repositories (e.g., NIH PubChem, BioSystems, NHANES, European Array Express)	All sources of policy-relevant data
New data types (Also uses the data from column to left)	QSAR, high-throughput *in vitro* screening assays, read- across, AOP development	High-content assays, medium throughput assays, knowledge-mined large data sets, AOP development	Molecular epidemiology, clinical and animal studies, AOP network development
Exposure paradigms of studies considered	*In vitro, in silico*	All relevant	All relevant
Metabolism in test systems	Some to none	Partial to intact	Intact
Incorporation of toxicokinetics	Reverse toxicokinetic models	Reverse toxicokinetics models, biomonitoring	Dosimetry and PK modeling, biomonitoring
Consideration of human variability and susceptibility	*In vitro* methods available	*In vitro* and *in vivo* methods available	*In vivo* methods available
Use of traditional *in vivo* data	*In vitro* assays anchored to pesticide registration and pharmaceutical data	None to limited; especially can be used in AOP development	New data types augment traditional; traditional data currently remain basis for assessment
Hazards	Nonspecific	Nonspecific to identified	Identified
Potency metrics	Relative rankings based on QSAR or HT toxicity values	Relative rankings and toxicity values	Risk distributions, cumulative & community risks
Likely strength of evidence linking exposure to effect	Suggestive	Suggestive to likely	Suggestive to known
Numbers of chemicals that can	10,000s	100s–1000s	100s

EPA-HQ-2018-000876_0073512

be assessed			
Time to conduct assessment	Hours–Days	Hours–Weeks	Days–Years

QSAR = quantitative structure activity relationship; HT = high throughput, EPA = U.S. Environmental Protection Agency, ACToR = Aggregated Computational Toxicology Resource (EPA), ToxCast = Toxicity Forecaster, NIH =

Figure 1. Three broad decision-context categories are shown across the top (white type), below which are the seven "fit-for-purpose" prototypes developed for this effort (black type). From left to right in Figure 1, the amount of traditional toxicological data available for assessment (e.g., *in vivo* rodent toxicity data, epidemiology data) and the confidence in the assessment conclusions decrease but the number of chemicals that can be evaluated increases markedly. PAHs = Polycyclic aromatic hydrocarbons; B[a]P = Benzo[a] pyrene.

National Institutes of Health, NHANES = National Health and Nutrition Examination Survey, AOP = adverse outcome pathway, PK = pharmacokinetic

EPA-HQ-2018-000876_0073513

Figure 2. 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. 2013). Reproduced with permission from *Environmental Health Perspectives*.