EPA Public Access

Author manuscript

Hum Ecol Risk Assess. Author manuscript; available in PMC 2020 November 06.

About author manuscripts | Submit a manuscript

Published in final edited form as:

Hum Ecol Risk Assess. 2019 November 6; 25: 1–24. doi:10.1080/10807039.2019.1615828.

Methods for evaluating variability in human health dose-response characterization

Daniel A. Axelrad[a], **R. Woodrow Setzer**[b,*], **Thomas F. Bateson**[c], **Michael DeVito**[d], **Rebecca C. Dzubow**[e], **Julie W. Fitzpatrick**[f,*], **Alicia M. Frame**[g,†], **Karen A. Hogan**[c,*], **Keith Houck**[b], **Michael Stewart**[b]

[a]Office of Policy, U.S. Environmental Protection Agency, Washington, DC, USA

[b]Office of Research and Development, U.S. Environmental Protection Agency, Research Triangle Park, NC, USA

[c]Office of Research and Development, U.S. Environmental Protection Agency, Washington, DC, USA

[d]National Institute of Environmental Health Sciences, National Toxicology Program, Research Triangle Park, NC, USA

[e]Office of Children's Health Protection, U.S. Environmental Protection Agency, Washington, DC, USA

[f]Office of the Science Advisor, U.S. Environmental Protection Agency, Washington, DC, USA

[g]Office of Land and Emergency Management, U.S. Environmental Protection Agency, Washington, DC, USA

Abstract

The Reference Dose (RfD) and Reference Concentration (RfC) are human health reference values (RfVs) representing exposure concentrations at or below which there is presumed to be little risk of adverse effects in the general human population. The 2009 National Research Council report *Science and Decisions* recommended redefining RfVs as "a risk-specific dose (for example, the dose associated with a 1 in 100,000 risk of a particular end point)." Distributions representing variability in human response to environmental contaminant exposures are critical for deriving risk-specific doses. Existing distributions estimating the extent of human toxicokinetic and toxicodynamic variability are based largely on controlled human exposure studies of pharmaceuticals. New data and methods have been developed that are designed to improve estimation of the quantitative variability in human response to environmental chemical exposures. Categories of research with potential to provide new database useful for developing updated

CONTACT Daniel A. Axelrad, axelrad.daniel@epa.gov, Office of Policy, U.S. Environmental Protection Agency, Washington, DC 20460, USA.

*Retired.

†Current address: 1: Neo4j, San Mateo CA 9440, USA.

Disclosure statement

No potential conflict of interest was reported by the authors.

The views expressed in this article are those of the authors and do not necessarily reflect the views or policies of the U.S. Environmental Protection Agency or the National Institute of Environmental Health Sciences.

EPA Author Manuscript

EPA Author Manuscript

EPA Author Manuscript

EPA Author Manuscript EPA Author Manuscript EPA Author Manuscript

human variability distributions include controlled human experiments, human epidemiology, animal models of genetic variability, *in vitro* estimates of toxicodynamic variability, and *in vitro-based* models of toxicokinetic variability. *In vitro* approaches, with further development including studies of different cell types and endpoints, and approaches to incorporate non-genetic sources of variability, appear to provide the greatest opportunity for substantial near-term advances.

Keywords

Human variability; dose-response; reference value; probabilistic risk; in vitro; risk assessment

Introduction

For characterizing noncancer risks from chronic exposures to toxic substances, U.S. Environmental Protection Agency (USEPA) risk assessments commonly use the oral Reference Dose (RfD) and inhalation Reference Concentration (RfC), route-specific exposure levels (oral or inhalation) at or below which it is presumed there is little risk of adverse effects in the general human population (U.S. EPA 2002, 2018). The main steps in deriving these human health reference values (RfVs) are 1) determination of a point of departure (POD), a dose at the lower end of the range at which adverse effects have been observed in a toxicological or epidemiological study of a chemical; and 2) adjustment of the POD by a set of factors that reflect relevant uncertainties or scientific judgments for a given data set in extrapolating to a suitable human population exposure level. This set of factors includes a default factor of 10 that has traditionally been applied in RfV derivation to represent human variability of response to a toxic substance exposure when chemical-specific population variability information is unavailable (U.S. EPA 2002).

The National Research Council's (NRC) report *Science and Decisions* (NRC 2009) observed that "The current formulation of the RfD is problematic because of its application as a determinant of risk vs. no risk of regulatory importance, and it lacks a quantitative description of the risk at different doses." The NRC recommended that USEPA adopt a "unified approach to dose-response assessment that will result in risk estimates for both cancer and noncancer end points" (NRC 2009). This approach would involve redefining RfVs for noncancer endpoints as "a risk-specific dose (e.g., the dose associated with a 1 in 100,000 risk of a particular end point), and the risk could be estimated at doses above and below the RfD" (NRC 2009). The NRC illustrated this approach through the application of statistical distributions to represent uncertainty and variability, in place of point estimate uncertainty factors currently used in RfV development.

The methods recommended by the NRC were extended by the International Programme on Chemical Safety (IPCS) (WHO 2014). Both the NRC and IPCS considered the application of distributions to depict the variability of human response to environmental contaminant exposures to be a critical element for deriving a risk-specific dose. That is, expressing the human variability factor as a distribution enables extrapolation from a dose estimated to result in, for example, 10% incidence of an effect to lower risk levels (e.g., a dose estimated to result in a 1% or 0.1% incidence). The NRC observed that "Default distributions that

EPA Author Manuscript

EPA Author Manuscript

EPA Author Manuscript

characterize [human] variability... will be needed as starting points that can be improved as the research advances" (NRC 2009). The IPCS advanced this approach with a set of default "preliminary distributions," derived primarily from a set of controlled human exposure studies of pharmaceuticals, along with some human epidemiological data (WHO 2014).

In recent years, new data and new methods have been developed that may enable improved estimation of the quantitative variability in human response to chemical exposures (Abdo et al. 2015b; Harrill and McAllister 2017; Wetmore et al. 2014; Zeise et al. 2013). To advance development of dose-response assessment concepts recommended in *Science and Decisions*, this paper reviews several types of data that could be utilized to improve upon the existing estimated distributions of human variability, including data from human, animal, and *in vitro* studies.

To provide context, we begin with a review the current approaches used by USEPA to address human variability in RfV derivation. We then further describe the recommendations of the NRC and the IPCS concerning human variability that have motivated this review. Finally, we consider several types of data that may be useful for further refining estimates of human variability.

Human variability in USEPA reference values

USEPA defines an RfV as "An estimate (with uncertainty spanning perhaps an order of magnitude) of an exposure to the human population (including sensitive subgroups) that is likely to be without an appreciable risk of deleterious effects during a lifetime" (U.S. EPA 2018). Derivation of an RfV begins with determination of a POD, which may be a no-observed-adverse-effect level (NOAEL), lowest-observed adverse effect level (LOAEL), or the lower confidence limit on a benchmark dose (BMDL). The RfV is then derived by dividing the POD by a set of uncertainty or adjustment factors that reflect limitations of the data underlying the POD. These factors are intended to account for up to five areas of extrapolation (interspecies differences, human variability, LOAEL-to-NOAEL, subchronic-to-chronic, and database uncertainty) (U.S. EPA 2002).

In USEPA RfV derivations, the human variability (or intraspecies uncertainty) factor accounts for "variations in susceptibility within the human population (interhuman variability) and the possibility (given a lack of relevant data) that the database available is not representative of the dose/exposure-response relationships in the subgroups of the human population that are most sensitive to the health hazards of the chemical being assessed" (U.S. EPA 2002).

In lieu of information to support an alternative, USEPA's traditional default value for the human variability factor has been 10-fold (U.S. EPA 2002). Lehman and Fitzhugh (1954) first applied an overall safety factor of 100 to the maximum safe dosage identified in long-term animal experiments, identifying interspecies and intraspecies variation as the components of this factor. They reported anecdotally that "a sick individual may be as much as 10 times more susceptible to toxic substances than an individual in good health," consistent with a 10-fold range of human response variability. Initial data analysis efforts to

EPA Author Manuscript

EPA Author Manuscript

EPA Author Manuscript

evaluate the extent of human variability supported the continued use of the 10-fold default factor for human variability (Dourson et al. 1996; Dourson and Stara 1983; Hattis et al. 1987).

Subsequent investigations evaluated variability among a variety of populations and lifestages, generally relying on human experimental data for pharmaceuticals. Many of these investigations divided overall human variability into toxicokinetic and toxicody-namic components. Toxicokinetic (TK) variability may be defined as "differences among people in the external dose required to produce a similar systemic internal dose," and results from variability in absorption, distribution, metabolism and excretion of a chemical. Toxicodynamic (TD) variability may be defined as "differences among people in the internal dose required to produce an effect of defined degree or severity" (WHO 2014). Renwick and Lazarus (1998) evaluated TK and TD variability for populations exposed to pharmaceuticals, including adult Caucasians, children, non-Caucasian ethnic groups, and those with polymorphic metabolisms. They concluded that a 10-fold factor from the population mean would cover 99.9% of the general population and noted that in some circumstances this 10-fold factor might be too low, such as certain situations involving very young children. Other studies have observed greater sensitivity among the elderly than among mature adults (Abdel-Megeed et al. 2001; Skowronski and Abdel-Rahman 2001). Hattis et al. (2002), also based on studying TK and TD among individuals exposed to pharmaceuticals, concluded that the use of the traditional 10-fold factor would be underprotective for children. Dorne (2007) evaluated human variability in hepatic and renal elimination and found that the TK portion of the factor "would not cover neonates, the elderly for most elimination routes and any subgroup of the population for compounds metabolized by polymorphic isozymes (such as CYP2C19 and CYP2D6)."

When sufficient chemical-specific data on human variability are available, USEPA has supported the development and use of data-derived extrapolation factors in place of the default 10-fold human variability factor (U.S. EPA 2014). For example, USEPA has used chemical-specific data in place of the default human TK variability factor in developing reference values for boron (U.S. EPA 2004), methylmercury (U.S. EPA 2001), dichloro-methane (U.S. EPA 2011a), and trichloroethylene (U.S. EPA 2011b). In particular, for the reference value for boron, USEPA used data on glomerular filtration rate in pregnant women to adjust the intraspecies uncertainty factor (U.S. EPA 2004).

Probabilistic methods using human variability distributions

Distributions of values to represent human variability have been explored most extensively in the context of proposed probabilistic methods for deriving reference values (Evans et al. 2001; Hattis et al. 2002; Swartout et al. 1998). The Hattis et al. (2002) publication, *A Straw Man Proposal for a Quantitative Definition of the RfD*, has been particularly influential and frequently cited. Hattis *et al.* proposed a quantitative definition for the RfD as "The daily dose rate that is expected (with 95% confidence) to produce less than 1/100,000 excess incidence over background of a minimally adverse response in a standard general population." They also outlined probabilistic methods by which an RfD satisfying the

EPA Author Manuscript

EPA Author Manuscript

EPA Author Manuscript

proposed quantitative definition could be derived. In brief, the approach for deriving such a value from an animal study of a quantal/dichotomous endpoint would involve:

1. Deriving an animal ED_{50} (dose associated with effects in the median member of the exposed group) with an uncertainty distribution, and applying adjustments as appropriate (e.g., for interspecies extrapolation) to derive an estimated human ED_{50} with an uncertainty distribution; and

2. Extrapolating from the estimated human ED50 to lower doses and response rates by "assuming a lognormal distribution of human susceptibilities, with uncertainty in the spread of human susceptibility drawn from an expanded database of human inter-individual variability observations."

Thus, the human variability distribution plays a critical role in probabilistic methods by enabling extrapolation from doses associated with a relatively high population response rate to estimate doses associated with lower population response rates. Hattis and colleagues also derived distributions of human TK and TD variability (Hattis et al. 2002; Hattis and Lynch 2007a) suitable for use in probabilistic modeling.

In *Science and Decisions*, the NRC applied these methods, including use of the Hattis human variability distributions, in its recommendations for a unified approach to dose-response assessment (NRC 2009). More recently, the IPCS (WHO 2014) elaborated on and extended the Hattis et al. (2002) approach, with specification of explicit procedures for addressing continuous endpoints, along with detail on implementation and standardized terminology and notation. The IPCS also reviewed the literature on derivation of human variability distributions and presented preliminary distributions for TK and TD.

The underlying concept from the literature concerning estimation of human variability in response to chemical exposures in general is to first obtain estimates of human variability from individual chemicals with appropriate data in existing studies, then characterizing the distribution of such estimates across those chemicals, and then applying that distribution to chemicals with inadequate data. The geometric standard deviation (GSD) is convenient for summarizing the population variability, where a larger GSD indicates greater variation in the population. For example, assume we have data on human TK variability for a subset of chemicals (while we lack such data for other chemicals), and the data for each chemical are summarized by a GSD. These chemical-specific TK GSDs themselves follow a distribution, as within a set of chemicals there will be lesser variability in response (smaller GSD) for some chemicals and greater variability (larger GSD) for other chemicals. Finally, assume that the distribution of GSDs for the subset of chemicals with TK variability data is reasonably representative of what we might expect as the distribution of GSDs for those chemicals that lack chemical-specific TK data. Thus, the median GSD among the chemicals with data can serve as a median estimate of the TK variability for chemicals lacking data, and the 95th percentile GSD of the chemicals with data can serve as the 95th percentile estimate of TK variability for chemicals lacking data. A corresponding approach can be taken to derive a distribution to represent TD variability for chemicals lacking TD data. These TK and TD distributions, based on subsets of chemicals with relevant data and

EPA Author Manuscript
EPA Author Manuscript
EPA Author Manuscript

assumed to be applicable to the universe of chemicals of interest lacking chemical-specific data, are referred to as “generic” human variability distributions.

The IPCS preliminary generic distributions for human variability (see Table 1) were derived from past work to determine such GSDs for multiple chemicals by Hattis and colleagues (Hattis 2013; Hattis et al. 2002; Hattis and Lynch 2007a). Most, but not all, GSDs used to develop these generic distributions were obtained from controlled human exposure studies, primarily in adults and primarily of pharmaceuticals.

The IPCS noted the following limitations of the preliminary human variability distributions. First, measurement errors in the underlying studies would likely result in overstatement of the extent of human variability. Second, the human subjects in the underlying studies, primarily healthy adults, are expected to be much less diverse than the general human population—potentially resulting in understatement of the extent of human variability. Third, the underlying studies were primarily studies of pharmaceuticals, as controlled exposure studies of most environmental contaminants are rarely if ever conducted, due to ethical constraints. Pharmaceuticals may differ from environmental chemicals in key properties relevant to variability in response (WHO 2014). Drugs generally have been developed against specific molecular targets and administered in doses that minimize interactions with other macromolecules. They also generally fall within a class of physicochemical properties that provide appropriate absorption/distribution/elimination properties to allow administration as oral drugs (WHO 2014). Environmental chemicals, on the other hand, are not typically developed for bioactivity (at least not for humans), may require biotransformation to toxic metabolites, and have a much broader range of physicochemical properties. Thus, there may be a much higher degree of potential interacting factors modulating toxicity *in vivo* and more opportunity for influence by human variability factors.

A further consideration is that the available estimates of human variability rely on data from relatively small sets of compounds. These data sets are quite small relative to the universe of controlled human exposure studies; one major reason for this is that studies were selected for inclusion only if they included individual observations as opposed to only summary statistics (Hattis and Lynch 2007a). In addition, Hattis completed assembly of his data set in 2007, and the date of publication for the included studies ranged from 1977 to 2000 (Hattis and Lynch 2007b).

Data representing TK and TD variability across larger and more diverse sets of compounds would increase confidence in the assumption that distributions derived from compounds with data are representative of what we would expect for compounds lacking data. It is also important to recognize that the identified limitations of these distributions are equally applicable to the studies described above that evaluated the suitability of the default 10-fold factor used to represent human variability in RfV derivation.

EPA Author Manuscript EPA Author Manuscript EPA Author Manuscript

Sources of information on human toxicokinetic and toxicodynamic variability

Given the limitations of the currently available TK and TD human variability distributions for probabilistic dose-response assessment, it is critical to explore additional data that could improve distribution estimates. A conceptually comprehensive characterization of inter-individual variability in susceptibility to environmental exposures, considering all of the various underlying determinants of variability, could require data: to identify potentially susceptible populations and lifestages; to define their unique characteristics based on physiology, behavior and co-exposures; and to describe how exposure can affect these groups differently from the general population. Variation in response among population groups could be due to many characteristics, including age, sex, genetics, race/ethnicity, lifestyle, preexisting disease status, microbiome, chemical co-exposures, and nutritional status. The combination of these factors can be instrumental in either promoting susceptibility to, or protecting from, certain disease outcomes, and their combined effect may depend on the lifestage when the exposures occur. Although characterizing all determinants of variability, accounting for their prevalence, magnitude and interactions, is not feasible, such an extensive characterization may not be necessary to develop reasonable estimates of the extent of variability in the population and how it differs across environmental contaminants. Instead, there are approaches available that may provide estimates of variability that integrate across many determinants.

The various types of data considered here, including data from controlled human exposure, epidemiologic, animal toxicology, and *in vitro* studies, may each provide a means for estimating human TK and/or TD variability overall, or for major components and determinants of that variability. They may, however, differ in the extent to which they reflect various sources of human variability (e.g., genetic vs. non-genetic sources).

Controlled human exposure studies

Controlled human exposure studies are experiments in which human participants are exposed to known doses of a given substance, with specified responses (e.g., measures of internal dose or some biological response) compared with those for a control exposure. Controlled human studies are often conducted for pharmaceuticals. Controlled human studies are also conducted for low doses of air pollutants (NASEM 2017), but are generally conducted infrequently for other environmental contaminants due to ethical constraints (NRC 2004).

Relative to other types of studies under consideration, the advantages of controlled human exposure studies are their experimental design with known exposure levels, and that they are conducted in humans (i.e., the species of interest for human health risk assessments). As described above, the existing estimates of human variability identified by the IPCS are largely based on controlled human exposure studies of pharmaceuticals (Hattis and Lynch 2007a; WHO 2014). An expanded and updated database of controlled human exposure studies could be compiled and used to develop updated distributions of TK and TD variability. Although such a database would likely have many of the same limitations as the

EPA Author Manuscript

EPA Author Manuscript

EPA Author Manuscript

existing estimates, it would likely also provide an improved basis for estimating TK and TD variability by incorporating a broader selection of studies, more recent observations, as well as more current research methods. If enough studies are found with sufficient information on chemicals that are not pharmaceuticals, there would be an opportunity to derive estimates that are more directly relevant to the universe of substances assessed by USEPA, and an opportunity to compare human variability data for pharmaceuticals to non-pharmaceuticals to assess the extent to which there are differences. In addition, a larger database might provide an opportunity to develop estimates specific to routes of exposure (e.g., inhalation, ingestion) or target organs, and perhaps an opportunity to assess the contributions to variability for specific factors (e.g., lifestyle, genetics), rather than only developing generic estimates for use in all risk assessments.

Although there would be many benefits to using data from controlled human exposure studies to inform TK and TD variability distributions, the use of data from these studies comes with a number of caveats. As noted in the IPCS comments on the Hattis data set, controlled human exposure studies typically involve the use of relatively healthy adults, and for ethical reasons likely do not include the most sensitive individuals. As a result, data from these studies are likely to underestimate the extent of variability in the general population. In addition, controlled human exposure studies are probably most informative for acute health effects given that these studies almost always consist of short-term exposures. While these are certainly limitations that would need to be considered, data obtained from these experiments should still be informative in developing updated variability distributions.

Epidemiological studies

Epidemiology is the study of the distribution and determinants of health in specified populations. Epidemiological studies of environmental exposures are usually observational, in that the research is conducted in populations as they are exposed to pollutants or other stressors on the job, in their communities and/or in their homes, rather than controlled experiments with specifically administered exposure levels. However, epidemiological studies are usually designed to estimate a mean effect in an exposed population, and less commonly assess the variability *(i.e.*, heterogeneity) of response in that population.

In general, epidemiological studies that characterize human variability in susceptibility primarily have focused on particular disease states and demographic variables such as age, sex, race/ethnicity, and socioeconomic status, rather than the full range of factors that may influence a variability of response within the population. For example, epidemiological studies have associated particulate air pollution with higher mortality rates in diabetics (Bateson and Schwartz 2004), and higher morbidity and mortality in the elderly (Bell et al. 2013; Simoni et al. 2015) relative to the effect of particulates in the general adult population. These large studies with detailed exposure data are important in identifying specific populations and lifestages with a greater risk of disease from air pollutants.

However, epidemiological studies that might assess variability of response in the population and characteristics associated with variability are often restricted to predefined individual, or group-level, characteristics available from central databases (e.g., census data) and are not designed to gain specific knowledge about human variation in response. Standard

EPA Author Manuscript EPA Author Manuscript EPA Author Manuscript

epidemiological study designs may limit characterization of human variability in exposure, often assuming, for example, that individuals with the same group-level exposure characterization are homogeneous with respect to exposure—even when there is a range of exposures within each exposure group (Loomis and Kromhout 2004). That is, epidemiological studies in general usually recruit a randomly-selected study population to avoid selection bias, and exposure groups and responses are usually described using measures of central tendency (Morimoto et al. 2003). In contrast, the assessment of interindividual variability generally involves analyzing which population groups are most and least responsive and thus expressing effects at the tails of the study response distribution where the data are sparser and effect estimates are less certain. Once a sensitive subgroup or suspected subgroup is identified (e.g., diabetics), additional studies may focus on that sensitive subgroup and how it differs from the broader population (Dubowsky et al. 2006).

Epidemiologic studies are more likely to detect any existing differences in response between different population groups when uncertainties from study limitations are minimized (Burns et al. 2014). Several recent advances in how epidemiological studies are designed and conducted are increasing their potential for informing quantitative characterization of human variability in response. These include improved exposure characterization (e.g., better measures of personal exposure—including biomarkers) in conjunction with improved characterization of genetic variability and other characterizations of personal susceptibility, and analyses using newer statistical approaches specifically aimed at characterizing variability in response (Burns et al. 2014; Schwartz et al. 2011).

For an epidemiological study to characterize human variability in response depends not only on study design characteristics (e.g., size and diversity of cohort, nature of exposure assessment, epidemiologic study design, information collected on potential confounders or effect modifiers), but also on the type of analysis applied. Regression models are usually used to test the influence of effect modifiers when studying differential susceptibility (Schwartz et al. 2011). With sufficient information on confounders or effect modifiers, standard regression models provide an estimated change in the mean response relative to a change in exposure and support characterization of response at different exposure levels. Other types of statistical analyses, however, might be more useful in describing variability of outcomes within a study population. Statistical techniques that could be used to estimate differences in response among individuals, or among groups within a cohort, include the following:

1. Estimation of individual effect estimates (slopes) for each subject in longitudinal studies can help determine whether subject-specific slopes vary in only a few susceptible individuals or in all subjects in the study (Naumova et al. 2001; Schwartz et al. 2005).

2. Case-crossover analysis, a version of a case-control study where the case also serves as the control (Maclure 1991), is used to estimate acute responses and can identify susceptibility factors by identifying effect modifiers because the same subject is exposed to the control and the experimental condition. For example, individuals who have diabetes are at increased risk to air pollution relative to those who have not (Bateson and Schwartz 2004). This differential risk could be examined by comparing the case-versus-control effect of air

EPA Author Manuscript

EPA Author Manuscript

EPA Author Manuscript

pollution in this population to the case-versus-control effect in the rest of the population (Schwartz et al. 2005).

3. Other models that have been suggested for studying human variability include: marginal structural models (causal models incorporating inverse probability of treatment weighting in order to account for the effect of time-dependent and previous treatment-or-exposure-dependent covariates), hierarchical mixed models (general or generalized linear models that account for group-specific correlations, such as the unique effects and interactions in certain populations or lifestages), and systems analysis (a method of examining higher-dimensional interactions) (Schwartz et al. 2011).

4. Quantile regression, which provides separate estimates of effects at various percentiles in the distribution and an unbiased estimate of the median, rather than the mean (Bind et al. 2015). Mean regression analyses may not capture associations that occur primarily in the tails of the outcome distribution, and quantile regression can be used to focus on individuals that have extreme responses to an exposure (more susceptible). Quantile regression applied to causal mediation analyses can identify the effect of an exposure on an outcome by a set of mediators, with different percentiles of the distribution of the outcome (Richiardi et al. 2013). Quantile causal mediation effects, the difference between the medians of two outcomes, have been suggested to study the more susceptible individuals in a cohort (Imai et al. 2010).

5. Genome-wide association studies have been suggested to identify effects on individual genotypes that also might be affected by environmental exposures. The use of this methodology to identify candidate genotypes can be statistically hampered by adjustments to the significance levels that account for multiple comparisons. Using interaction terms in gene-wide association/interaction studies to evaluate specific pathways can be a more statistically powerful approach to detect gene-environment interactions (Schwartz 2015). The interaction terms—which are a multiplicative factor of the gene (or a set of genes in a specific pathway) and the environmental exposure—can be used to quantify the differences in exposure effect sizes between the genotypically susceptible and non-susceptible populations.

The advantage of epidemiological studies for characterizing human variability is that they represent actual exposure patterns and circumstances in humans. These studies can provide integrated measures of variability that may represent the influence of genetic and non-genetic factors, including, for example, the impacts of exposures to other environmental contaminants, nutrition and stress. However, as noted above, various aspects of study design may limit the ability to detect human variability in response in any particular study (e.g., assuming all individuals within a certain radius of a pollution monitor have the same level of exposure). Furthermore, given the lack of similarly designed epidemiological studies of the same exposures and the same outcomes measured with the same level of specificity, it may be difficult to obtain comparable estimates of human variability across a number of chemicals to inform generic human variability distributions. There is, however, the potential for a wealth of information on variability (susceptibility) for the criteria air pollutants:

EPA Author Manuscript

EPA Author Manuscript

EPA Author Manuscript

particulate matter (U.S. EPA 2009), ozone (U.S. EPA 2013b), sulfur dioxide (U.S. EPA 2017), nitrogen dioxide (U.S. EPA 2016a), and lead (U.S. EPA 2013a).

Animal models of genetic variability

Laboratory animal toxicology experiments have not traditionally been a source of information on the extent of human variability. However, several recent research efforts have developed population-based rodent models of genetic diversity. While direct quantitative inferences about human variability from rodent variability is problematic, rodent models may still provide mechanistic insight into the impact of human genetic diversity on the variability in response to chemical exposures

In animal-based toxicological research, there is a balance between the types of animal models used: those that are efficient, reproducible and sensitive, versus those that provide insight into varying population sensitivities. The first type of model leads to the use of inbred animals, while the second requires models having much broader genetic diversity. The use of inbred animals became popular because of their fixed and known genetics and because replicate animals with known phenotypic patterns of biology and disease are readily generated. Because inbred mice share the same genome, response to chemical exposure is much less variable, allowing fewer animals to achieve an appropriately powered study for a given endpoint. Rats and mice more generally, however, have broad genetic diversity, and genetic polymorphisms can significantly affect the responses to chemical toxicity (Rusyn et al. 2010). Studies using animal models of population variability with broader genetic diversity require more animals (*i.e*., are less efficient) because, for most endpoints of concern, there is larger variability in the controls (Harrill et al. 2017).

Approaches to examining variability in response across animal strains have focused on the mouse, probably due to the history of mouse breeding and the hundreds of mouse strains commercially available. Mouse panels are groups of inbred animals or recombinant crosses for which the genotype and phenotypes are well characterized and which represent a broad-based genetic diversity. One example is a series of studies that examined toxicity response to acetaminophen of priority strains in the Mouse Phenome Project. Harrill *et al.* (2009a) were able to identify biomarkers of response and identify a polymorphism in *CD44* that contributes to increased susceptibility to acetaminophen toxicity, and Harrill *et al.* (2009b) demonstrated that this polymorphism is also involved in human susceptibility.

The Collaborative Cross (CC) was designed to address the limitations of using small sets of inbred lines as genetic reference populations. The CC line was developed starting with eight founder strains representing both classical inbred strains and wild-derived strains. These animals were randomly bred for three generations and then inbred starting with the fourth generation (Threadgill and Churchill 2012; Threadgill et al. 2011). While this breeding scheme initially led to over 400 unique mouse lines, most of these lines became extinct and presently only about 74 CC lines remain available (Shorter et al. 2017; UNC Systems Genetics 2018). The eight founder strains have more genetic diversity than existing recombinant inbred panels, their genetic variation is more uniformly distributed throughout the genome compared with other inbred panels, and all founder strains in the CC have been sequenced (Threadgill and Churchill 2012; Threadgill et al. 2011). These mice provide a

EPA Author Manuscript EPA Author Manuscript EPA Author Manuscript

more complete randomization of genotypes than is available in other mouse reference populations, allowing for greater power and accuracy in genetic mapping and phenotype correlation analyses (Bogue et al. 2015).

While the CC mice have not yet been used to evaluate chemical toxicity, the CC founder strains have been used to evaluate inter-strain variability in responsiveness to chemical exposure because they represent the genetic diversity present in the three main subspecies of house mouse and their genomes have been fully sequenced (Koturbash et al. 2011). Yoo and colleagues have evaluated altered sensitivity to trichloroethylene using the CC founder strains and concluded that the difference in hepatotoxicity and nephrotoxicity could be explained in large part by observed inter-strain TK variability, specifically variations in TCE metabolite concentrations in the target tissues (Yoo et al. 2015a, 2015b). A study using the CC founder strains found that pulmonary toxicity of quantum dots was dependent on inter-strain differences in inflammatory response (Scoville et al. 2015).

A second new mouse model is the Diversity Outbred (DO), in which each mouse has a unique genotype (Church et al. 2015; French et al. 2015). The animals were derived from 144 partially inbred CC mice between generations F4 and F12 in the inbreeding scheme. Genetic variation is uniformly distributed with multiple allelic variants, making the DO a powerful tool for correlating genotype with phenotype and for high-resolution genetic mapping.

DO mice have been used to evaluate differential sensitivity in two studies. French et al. (2015) evaluated benzene-induced genotoxicity, as measured by micronucleated reticulocytes, in a 28-day inhalation study using the DO. This study used three dose levels and a control with two blocks of 75 mice/treatment group. French *et al.* derived a benchmark dose (BMD) for benzene in the DO mice that was 10 times lower than the BMD for the same endpoint from a study using male B6C3F1 mice.

The DO study of epigallocatechin gallate, a constituent of green tea, demonstrates both the possibilities and limitations of this model. Church et al. (2015) identified a subset of DO mice that was extremely sensitive to the hepatotoxic effects of epigallocatechin gallate and found a 9Mb region on chromosome 4 that contained 49 genes that might be related to enhanced sensitivity to the chemical. The authors translated this finding to human clinical studies and found that three of the proposed genes were associated with increased risk of hepatotoxicity in humans (Church et al. 2015). Because only two dose levels were used due to cost, the study provides no insight into how much more sensitive the susceptible population is with respect to dose. Therefore, although this study did identify a susceptible subpopulation based on genetic background, it provides no information on the magnitude of variability.

These new rodent models should provide improved approaches to understand and predict human population variability in response to chemical exposures as compared to use of inbred strains. The studies by French et al. (2015) and Church et al. (2015) demonstrate the models' strength in identifying genetic polymorphisms that influence chemical toxicity. Directly translating identified polymorphisms from these rodent models to humans is

EPA Author Manuscript EPA Author Manuscript EPA Author Manuscript

challenging because not all mouse genes have human orthologues (French et al. 2015). Experience using these models to evaluate variability in chemical sensitivity is limited, and these studies have focused on identifying genetic polymorphisms involved in altered susceptibility. Another challenge is the large number of animals necessary to identify genetic polymorphisms in these experiments. Churchill et al. (2012) estimated that, depending on the phenotype studied, between 200 and 800 mice per treatment group are required to identify any polymorphism related to the variance in response with sufficient confidence. As it stands, the use of mouse models to efficiently identify polymorphisms and quantify population variability in response to chemical exposure requires further optimization of study designs.

In addition, the DO model may not work well for certain outcomes of high interest; for example, to have the same statistical power of a guideline 90-day subchronic study using 10 B6C3F1 mice per dose group, it would require 27 DO mice per dose group to detect differences in body weight changes and 150 DO mice per dose group to detect changes in sperm count (Harrill et al. 2017). Changes in hormones may require over a 1000 DO mice per dose group to attain the same statistical power as 10 B6C3F1 mice per dose group (Harrill et al. 2017). For developmental toxicity studies, the limited historical data in breeding and developmental milestones in the DO mice make an appropriately powered study design uncertain. These mouse models also can account only for variability due to genetic differences, and do not represent other sources of human TK and TD variability, such as diversity in background exposure to other environmental contaminants.

To provide support for estimation of generic distributions of human variability applicable in situations where chemical-specific data are lacking, studies of response variability across multiple chemicals will likely be needed—an objective that will be resource intensive for both cost and time. In the interim, targeted studies on chemicals of interest may be useful for informing human genetic variability distributions in assessments of those specific chemicals, and ultimately a compilation of many such studies could be used to develop generic distributions. The use of these rodent models of genetic diversity in understanding population variability to chemical exposure will require concerted collaboration among the experimenters who generate the data and risk assessors who use these findings.

In vitro-*based estimates of toxicodynamic variability*

In vitro toxicity testing has been touted as the future of toxicology: a tool for high-throughput, relatively low-cost, and quick toxicity assessment. The use of modern *in vitro* systems allows testing of chemicals in cell-based or cell-free assays with sufficient throughput to test hundreds to thousands of chemicals against specific molecular endpoints quickly and inexpensively. Such efficiency is beginning to be applied towards informing estimates of variability in human response.

In vitro assays used for toxicity testing have traditionally used immortalized, transformed cell lines maintained in culture (Carmichael et al. 1987; Vichai and Kirtikara 2006). These genetically homogeneous cell lines were derived from individuals without regard to genetic background, sex, age, or other parameters known to affect human variability. Furthermore, many such cell lines have undergone an adaptation to growth under highly artificial *in vitro*

EPA Author Manuscript EPA Author Manuscript EPA Author Manuscript

conditions with both genetic and epigenetic changes occurring relative to the originally isolated cell (Hartung and Daston 2009). These changes and the lack of genetic characterization likely make these cell lines poor choices for assessing variability in the human population.

Several *in vitro* studies have been conducted with the specific goal of estimating toxicodynamic variability in response due to human genetic differences using more relevant cell types. Choy et al. (2008) examined a panel of 269 human lymphoblastoid cell lines (LCLs) from the International HapMap Project to determine variability in response to 7 anticancer agents with different mechanisms of action. The LCLs were developed using a standardized immortalization protocol making comparison across individual lines more robust, and also provided extensive publicly available genotypic sequencing data. Although these authors demonstrated variability in response to the drugs, they were unable to demonstrate linkages with genetic variation and could not rule out spurious associations. O'Shea et al. (2011) also used HapMap lines and found that responses varied for some, but not all, of 14 chemicals tested in 85 cell lines for cell viability and in 83 cell lines for caspase activity. They too could not completely rule out influences beyond genetic differences and had limited success mapping SNPs to variable responses. A follow-up study extended the testing to 240 chemicals and 81 cell lines, again measuring cytotoxicity and caspase activity but with more extensive concentration-response testing (Lock et al. 2012). Toxicodynamic variability across cell lines for both assay endpoints ranged up to three orders of magnitude for some chemicals; however, for the majority of the chemicals, TD variability across cell lines fell within a 10-fold range in difference between most sensitive and most resistant. While the larger observed differences were greater than the 10-fold default value currently used for human variability in risk assessment, more research is needed to ensure these were not experimental outliers and that the response is relevant to human genetic variability.

Abdo *et al.* (2015b) measured induced cytotoxicity, using 179 chemicals and 1086 human LCLs from the 1000 Genomes Project. These cell lines represented nine populations from five continents. The authors reported a median difference of 3-fold between the most sensitive 1% of cell lines compared with the median value across all cell lines, depending on the chemical. Extremes up to a 30-fold difference between the most sensitive and the median were noted using shrunken estimates that accounted for sample testing variability. Limitations of the study include replicates performed on the same day and no assessment of confounding factors such as individual cell-line growth rate. Nevertheless, the large study size study provides a useful overview of the general range of variability expected, although outliers deserve more careful study.

In a related study, Abdo *et al.* (2015a) measured variability in response to exposure to two defined pesticide mixtures across 146 human LCLs representing four ancestrally and geographically diverse populations. They found an inter-individual TD variability ratio between the median and 95th percentile cell lines of around 3-fold for each mixture. For further applications of this general approach, Chiu et al. (2017) noted the impracticality of screening thousands of cell lines against individual toxicity targets to derive TD ranges. They proposed a tiered experimental strategy of fewer, representative cell lines coupled with

EPA Author Manuscript EPA Author Manuscript EPA Author Manuscript

a Bayesian analysis method to flag chemicals with greater-than-default variability for more extensive characterization for population-wide responses.

A caveat to the approach of these human-based LCL studies is the reliance on only a single type of endpoint, cell viability (along with caspase activity, a related endpoint) for determining the range of variability in response. While this covers a broad range of potential toxicity mechanisms of action, there are many other important endpoints that would not necessarily affect cell viability and, thus, the effect of genetic variation would not be observed and the range of variability may differ. Furthermore, the existing research using cell line panels derived from diverse human populations has relied on the LCLs of the HapMap and 1000 Genomes projects rather than on primary cells derived from other tissues. Hence, there may be a lack of sensitivity to detecting organ-or other tissue-specific toxicities.

One technique that may provide somewhat more relevant information is the use of induced pluripotent stem cells (iPSCs). Derived from relatively easily accessible adult somatic cells, iPSCs could be sampled for a diverse human population (Takahashi et al. 2007; Yu et al. 2007). Various induction protocols are available that can reprogram these cells in to liver, brain, heart, and pancreas cell lineages (Brennand et al. 2011; Huangfu et al. 2008; Si-Tayeb et al. 2010; Zhang et al. 2009). Induction of iPSC liver cells may be particularly useful as these cells have xenobiotic metabolism activity that could be greatly influenced by genetic variability and, hence, diversity in the types and quantities of metabolites generated as in human liver (Evans and Relling 1999; Mann 2015). The iPSC-derived liver cells display a partial fetal phenotype, however, and are not yet fully capable of representing adult donor variability.

In addition to providing information on variability in response due to genetic differences, *in vitro* approaches could be considered to gain a better understanding of chemical response variability due to sex, lifestage, and non-chemical-stressors. With respect to sex, responses between male-derived versus female-derived cells could be compared. However, transformed, immortalized cell lines could have lost true differences in sex-linked patterns of response and so should be carefully genotyped and phenotyped before use for such purposes (Park et al. 2006). Use of primary cells or perhaps iPSCs, combined with a focus on endpoints expected to have other potential sex-linked responses, could prove useful. A major limitation to these models is lack of a true endocrine regulation function with current *in vitro* approaches, which may result in an underestimate of differences by sex.

Incorporating different lifestages is a major challenge for *in vitro* approaches. Early development is a special case where much effort has been devoted to developing useful models of chemical sensitivity, typically using human or mouse embryonic stem cells (Kameoka et al. 2014; Scholz et al. 1999; West et al. 2010). Alternatively, model organisms such as zebrafish have been employed during their development phase (5 days for zebrafish) (Hill et al. 2005). By comparing lowest effect levels in these systems to those from non-embryonic developmental models, it may be possible to identify chemicals of most concern to the fetus and begin to understand the range of differences in sensitivity. Other lifestages, such as childhood and old age, are not readily distinguishable with *in vitro* systems like cell

EPA Author Manuscript EPA Author Manuscript EPA Author Manuscript

culture models. Lapasset et al. (2011) found human iPSCs generated from senescent and centenarian cells have reset much of their cellular physiology to earlier stages and are indistinguishable from human embryonic stem cells. In contrast, Kang et al. (2016) showed mitochondrial DNA defects increased with age and were maintained in iPSCs derived from older donors. While it may be technically possible to develop cell models derived from specific subsets of the population with respect to lifestage through iPSC technology or use of primary cells, extensive validation will be required (Studer et al. 2015).

Finally, the effects of exposure to multiple chemical and non-chemical stressors in combination are known to influence response variability. For example, various types of stress, such as inflammation or poor nutrition; co-exposure to other environmental chemicals; or influence of the microbiome all vary between individuals and can alter the TK or TD response to the chemical of interest (NASEM 2018; Sexton and Hattis 2007; Van de Wiele et al. 2005). The effect of chemical mixtures in particular might be conducive to investigation with *in vitro* methods. Although measuring interactions of chemicals in all combinations that humans are exposed to is a large-scale problem that even high-throughput approaches have not yet addressed, alternative strategies might be valuable. For example, determining the effects of single chemicals in the context of their presence in complex environmentally relevant mixtures could identify those chemicals most likely to have significant interactions with other commonly occurring chemicals. This strategy could help identify chemicals for which the response at the cellular level is most likely to vary with the presence of other chemicals.

In vitro toxicity testing approaches can produce estimates of TD variability by testing hundreds to thousands of chemicals quickly and inexpensively in large numbers of cell lines representative of human genetic diversity. Toxicity responses in cell-based assays can show variability both across cell lines (due to underlying genetic and epigenetic variability) and within cells from the same line (due to variability in assay design and stochastic differences in response among individual cells). The cell types used in these studies to date may not be broadly representative, but advances in technology (new types of cell lines) should improve the relevance and applicability of *in vitro* data to estimation of human variability.

Although *in vitro* systems might be able to model many aspects of human individual variability manifested at the molecular or cellular level adequately, they cannot fully reflect systems-level effects at tissue and higher levels of biological integration. In particular, *in vitro* systems are not able to capture the compensatory biological responses that may occur in the whole organism, as well as the interactions (*e.g.*, limited compensation) with background biological processes such as the initiation of disease. Furthermore, *in vitro* systems to date have had limited capability to represent non-genetic factors (such as co-exposure to other environmental chemicals) that also influence variability.

Assessment of kinetic variability with generic toxicokinetic models and in vitro *methods*

Another application of *in vitro* data for estimation of human variability is in the use of generic TK models to estimate variability of internal dose resulting from a given administered dose (Ring et al. 2017; Rotroff et al. 2010; Wetmore et al. 2014). Population variability in factors that affect TK of a chemical, such as differences in physiology and

EPA Author Manuscript EPA Author Manuscript EPA Author Manuscript

metabolism, mean that for a homogeneous exposure, the relevant internal dose can be expected to vary across members of a population. *In vitro* testing can generate data used to represent key parameters (and their variability) in generic TK models.

Directly relating important aspects of TK variability that can be measured, such as activity of clearance enzymes, rates of absorption and physiological parameters such as body mass index, to variability of internal dose is often not straightforward (Lipscomb et al. 2003), but requires the application of TK models. TK models translate an oral, inhalation, or dermal dose into a target tissue dose (or a blood concentration, which often serves as a surrogate for target tissue dose). They are constructed as systems of differential equations describing the behavior of chemical concentrations in connected compartments. Such models range in complexity from simple one- and two-compartment models that primarily predict the time course of blood concentrations, to much more complex, physiologically based toxicokinetic (PBTK) models. In PBTK models, the compartments correspond to individual tissues connected by flowing arterial and venous blood and the models predict concentrations in all compartments included (Thompson et al. 2008). In one-compartment models, the parameters have intuitive interpretations: a parameter quantifying the volume of the body in which the chemical is assumed to be evenly distributed, a parameter that governs the rate at which the chemical is cleared from the body through metabolism and glomerular filtration, and a parameter that quantifies absorption unless dosing is intravenous. In contrast, parameters of PBTK models reflect more detailed physiological characteristics, such as blood flows and tissue volumes, and chemical-specific characteristics such as constants describing the partitioning of the chemical between blood and tissues or metabolic rates; parameters governing the rate of absorption across the possible routes of entry (*i.e.*, oral, dermal, or inhalation); parameters that quantify the degree of binding of chemicals to proteins; and parameters that quantify active transport of the chemical across various barriers (Thompson et al. 2008).

Conventionally, PBTK models have been constructed for individual chemicals or groups of structurally and functionally similar chemicals, using *in vivo* animal data comprising concentration time courses in multiple tissues at multiple doses. Such models can be extrapolated to estimate internal dose in humans, and can incorporate chemical-specific aspects of absorption and metabolism (Tan et al. 2018; Thompson et al. 2008). However, they are expensive and time-consuming to construct. A variety of generic TK models has been produced that are consistent with known physiology and, while typically less accurate than models built specifically for particular chemicals, are still useful for characterizing variability of tissue dose resulting from a given administered dose.

Generic TK models are used for chemicals for which little or no *in vivo* TK data are available. Examples of such models include Simcyp (see www.certara.com), PopGen (McNally et al. 2014), and the R package httk (https://cran.r-project.org/web/packages/httk/index.html) (Ring et al. 2017). Rotroff et al. (2010) and Wetmore et al. (2012, 2014) describe use of Simcyp to convert *in vitro* concentrations in ToxCast™ (Kavlock et al. 2012) assays to an equivalent steady-state dose rate. Estimates of population variability in TK parameters were used as inputs in a model to predict the likely range of such steady-state doses. A useful metric in this computation is the concentration at steady-state (C_{ss}), the

EPA Author Manuscript EPA Author Manuscript EPA Author Manuscript

plasma concentration that would result from a repeat daily dose of, for example, 1 μg/kg/day. Wetmore et al. (2014) estimated the C_{ss} for various demographic groups using measured *in vitro* clearance rates for nine chemicals, catalyzed by 13 human cytochrome P-450 enzymes and five 5′-diphospho-glucuronosyltransferase enzymes. By combining the chemical-specific *in vitro* clearance data with information about the variability of enzyme activity and physiology among demographic groups and among individuals within groups, the authors were able to calculate C_{ss} ratios between the median healthy adult and median and 95th percentile of the most sensitive demographic group considered. The ratios of overall median to sensitive group median across nine chemicals ranged from 1.3 to 4.3, and the ratio of overall median to sensitive group 95th percentile ranged from 3.1 to 13.1. For seven of the nine chemicals, the 0 to 0.5 year old age group was the most sensitive, while patients with renal failure were projected to be the most sensitive for the other two chemicals. This approach illustrates how the use of chemical-specific *in vitro* assays, along with generic models and information that is not chemical-specific, can be used to estimate the magnitude of human TK variability of environmental chemicals.

Conclusions

Reports from the NRC and the IPCS have recommended methods for estimating a risk-specific dose and described the advantages of this type of estimate relative to traditional RfVs that are not associated with any specified level of risk (NRC 2009; WHO 2014). Advantages of a risk-specific dose include: 1) the underlying methods could be used to estimate risk at any level of exposure to a substance, and 2) the methods provide an estimate of statistical confidence in those risk estimates. Estimates of human variability in response to chemical exposures, expressed as probabilistic distributions, are a critical input to these methods. The preliminary distributions for human TK and TD variability identified by the IPCS are based on data compiled from an array of historical studies, and could be used in risk assessments with perhaps as much confidence as standard default uncertainty factors. However, there is potential for improving on those preliminary distributions with further mining of data from existing experimental and epidemiologic studies in humans, as well as by incorporating newly emerging data from animal and *in vitro* studies.

Each type of data discussed in this paper offers different advantages and limitations for development of human variability distributions. Controlled human exposure studies offer the advantages of experimental design and human subjects; key limitations of these studies are that the study populations are generally small and homogeneous, not representative of the diversity of the human population, and the ethical constraints on conducting these studies with most environmental contaminants. Human epidemiological studies, especially those conducted in large general population samples, have the potential to provide the most direct and relevant representation of human diversity, incorporating genetic and non-genetic sources of variability; but their ability to discern variability may be limited by potential confounding, exposure misclassification and lack of data for key variables that are sources of variability. Toxicological studies conducted in diverse strains of mice have potential as a model of human variability associated with genetic differences; but experience with these models to date is quite limited (only two substances tested), and they may not represent important non-genetic sources of variability. *In vitro* studies of the influence of genetic

EPA Author Manuscript EPA Author Manuscript EPA Author Manuscript

diversity on TD variability in response have been conducted with large numbers of environmentally-relevant chemicals and in large numbers of cell lines; but endpoints and cell types used have been limited (primarily cytotoxicity in lymphoblastoid cell lines) and thus, may not be broadly representative. *In vitro* data, in conjunction with generic TK models, have also been used to characterize TK variability, but in only a limited number of substances to date.

Each approach described in this paper has varying potential for providing substantial new data for quantifying human variability over the next several years; further development of *in vitro* approaches, including studies of different cell types and endpoints, as well as approaches to incorporate non-genetic sources of variability, appears to provide the greatest opportunity for substantial near-term advances. *In vitro* studies can be conducted relatively quickly; the science is advancing rapidly; high-throughput *in vitro* studies routinely generate data for large numbers of environmental chemicals, and are able to incorporate extensive genetic variability in studies using existing cell lines. Ultimately, however, the best approach to estimating human variability distributions may draw on information from multiple areas of research, drawing on the strengths of one study design to overcome the limitations of another, and integrating results to incorporate both genetic and non-genetic sources of variability. In addition, there may be ways in which the different types of research can be combined; for example, *in vitro* study designs might be applied to cells obtained from genetically diverse mice, and *in vitro* studies may inform selection of priorities and approaches for human and animal studies.

Despite the focus here on deriving probabilistic estimates of human variability from existing data, it is also important to note that the approaches and data types considered for developing updated human variability distributions could also directly support data-derived human variability factors for RfVs, depending on the needs of the analysis or assessment under consideration. One outcome could be the development of refined categories of human variability estimates (e.g., target-organ-specific human variability estimates), with the result being greater flexibility for the risk assessor and analyses that better estimate the risk associated with chemical exposure. Thus, within these data types exists the potential for a suite of approaches by which to improve variability and uncertainty characterization, through the development of probabilistic or improved point estimates of human variability.

Researchers conducting studies focused on a particular substance (or mixtures of substances) in any of the disciplines noted above may not usually consider the use of their data in constructing distributions of human variability. Awareness of these applications, and communication with risk assessors, may help researchers identify opportunities to provide useful information with relatively minor enhancements to their study designs, statistical techniques and results reporting. The recent increase in requirements from journals and research funders for open access to data (AAAS 2018; U.S. EPA 2016b), along with data already in public repositories (NIH 2018), may also lead to a rapid increase in opportunities for secondary analysis of research data to assess human variability across numerous environmental contaminants in multiple types of studies, accounting for multiple sources of variability.

EPA Author Manuscript EPA Author Manuscript EPA Author Manuscript

Acknowledgments

The authors appreciate the valuable input of Weihsueh Chiu and Anna Lowit in developing the plan for this review. Andrea Pfahles-Hutchens contributed to the preparation of this paper. We appreciate the helpful input of numerous USEPA colleagues, including members of USEPA's Risk Assessment Forum. Development of this review was informed by the presentations and discussion at the National Academy of Sciences Emerging Science for Environmental Health Decisions Workshop, Interindividual Variability: New Ways to Study and Implications for Decision-Making, held on September 30 and October 1, 2015.

References

AAAS. 2018 Science Journals: editorial policy. Available from: http://www.sciencemag.org/authors/science-journals-editorial-policies [accessed 11 June 2018]

Abdel-Megeed MA, Suh DH, and Abdel-Rahman MS. 2001 Intra-species extrapolation in risk assessment of different classes of antimicrobials. Hum Ecol Risk Assess 7:15–36

Abdo N, Wetmore BA, Chappell GA, et al. 2015 In vitro screening for population variability in toxicity of pesticide-containing mixtures. Environ Int 85:147–55 [PubMed: 26386728]

Abdo N, Xia M, Brown CC, et al. 2015 Population-based in vitro hazard and concentration-response assessment of chemicals: The 1000 Genomes high-throughput screening study. Environ Health Perspect 123:458–66 [PubMed: 25622337]

Bateson TF, and Schwartz J. 2004 Who is sensitive to the effects of particulate air pollution on mortality? A case-crossover analysis of effect modifiers. Epidemiology 15:143–9 [PubMed: 15127905]

Bell ML, Zanobetti A, and Dominici F. 2013 Evidence on vulnerability and susceptibility to health risks associated with short-term exposure to particulate matter: A systematic review and meta-analysis. Am J Epidemiol 178:865–76 [PubMed: 23887042]

Bind MA, Coull BA, Peters A, et al. 2015 Beyond the mean: Quantile regression to explore the association of air pollution with gene-specific methylation in the Normative Aging Study. Environ Health Perspect 123:759–65. [PubMed: 25769179]

Bogue MA, Churchill GA, and Chesler EJ. 2015 Collaborative cross and diversity outbred data resources in the Mouse Phenome Database. Mamm Genome 26:511–20. [PubMed: 26286858]

Brennand KJ, Simone A, Jou J, et al. 2011 Modelling schizophrenia using human induced pluripotent stem cells. Nature 473:221–5. [PubMed: 21490598]

Burns CJ, Wright JM, Pierson JB, et al. 2014 Evaluating uncertainty to strengthen epidemiologic data for use in human health risk assessments. Environ Health Perspect 122:1160–5. [PubMed: 25079138]

Carmichael J, DeGraff WG, Gazdar AF, et al. 1987 Evaluation of a tetrazolium-based semiautomated colorimetric assay: assessment of chemosensitivity testing. Cancer Res 47:936–42. [PubMed: 3802100]

Chiu WA, Wright FA, and Rusyn I. 2017 A tiered, Bayesian approach to estimating of population variability for regulatory decision-making. ALTEX 34:377–88 [PubMed: 27960008]

Choy E, Yelensky R, Bonakdar S, et al. 2008 Genetic analysis of human traits in vitro: Drug response and gene expression in lymphoblastoid cell lines. PLoS Genet 4:e1000287. [PubMed: 19043577]

Church RJ, Gatti DM, Urban TJ, et al. 2015 Sensitivity to hepatotoxicity due to epigallocatechin gallate is affected by genetic background in diversity outbred mice. Food Chem Toxicol 76:19–26. [PubMed: 25446466]

Churchill G, Gatti D, Munger S, et al. 2012 The diversity outbred mouse population. Mamm Genome 23:713–8. [PubMed: 22892839]

Dorne JL. 2007 Human variability in hepatic and renal elimination: Implications for risk assessment. J Appl Toxicol 27:411–20 [PubMed: 17497760]

Dourson ML, Felter SP, and Robinson D. 1996 Evolution of science-based uncertainty factors in noncancer risk assessment. Regul Toxicol Pharmacol 24:108–20 [PubMed: 8933624]

Dourson ML, and Stara JF. 1983 Regulatory history and experimental support of uncertainty (safety) factors. Regul Toxicol Pharmacol 3:224–38 [PubMed: 6356243]

EPA Author Manuscript EPA Author Manuscript EPA Author Manuscript

Dubowsky SD, Suh H, Schwartz J, et al. 2006 Diabetes, obesity, and hypertension may enhance associations between air pollution and markers of systemic inflammation. Environ Health Perspect 114:992–8. [PubMed: 16835049]

Evans JS, Rhomberg LR, Williams PL, et al. 2001 Reproductive and developmental risks from ethylene oxide: A probabilistic characterization of possible regulatory thresholds. Risk Anal 21:697–718. [PubMed: 11726021]

Evans WE, and Relling MV. 1999 Pharmacogenomics: translating functional genomics into rational therapeutics. Science 286:487–91 [PubMed: 10521338]

French JE, Gatti DM, Morgan DL, et al. 2015 Diversity outbred mice identify population-based exposure thresholds and genetic factors that influence benzene-induced genotoxicity. Environ Health Perspect 123:237–45. [PubMed: 25376053]

Harrill A, Borghoff S, Zorrilla L, et al. 2017 NTP research report on baseline characteristics of diversity outbred (J:DO) mice relevant to toxicology studies NTP Research Report 6, Research Triangle Park, NC.

Harrill AH, and McAllister KA. 2017 New rodent population models may inform human health risk assessment and identification of genetic susceptibility to environmental exposures. Environ Health Perspect 125:086002 [PubMed: 28886592]

Harrill AH, Ross PK, Gatti DM, et al. 2009 Population based discovery of toxicogenomics biomarkers for hepatotoxicity using a laboratory strain diversity panel. Toxicol Sci 110:235–43 [PubMed: 19420014]

Harrill AH, Watkins PB, Su S, et al. 2009 Mouse population-guided resequencing reveals that variants in CD44 contribute to acetaminophen-induced liver injury in humans. Genome Res 19:1507–15 [PubMed: 19416960]

Hartung T, and Daston G. 2009 Are in vitro tests suitable for regulatory use? Toxicol Sci 111:233–7 [PubMed: 19617452]

Hattis D 2013 Human interindividual variability in parameters related to susceptibility for toxic effects [website]. Available from: http://www2.clarku.edu/faculty/dhattis [accessed 30 April 2014]

Hattis D, Baird S, and Goble R. 2002 A straw man proposal for a quantitative definition of the RfD. Drug Chem Toxicol 25:403–36 [PubMed: 12378950]

Hattis D, Erdreich L, and Ballew M. 1987 Human variability in susceptibility to toxic chemicals—A preliminary analysis of pharmacokinetic data from normal volunteers. Risk Anal 7:415–26 [PubMed: 3444929]

Hattis D, and Lynch MK. 2007a Empirically observed distributions of pharmacokinetic and pharmacodynamic variability in humans—Implications for the derivation of single-point component uncertainty factors providing equivalent protection as existing reference doses In: Lipscomb JC and Ohanian EV (eds), Toxicokinetics in Risk Assessment, pp. 69–93. Informa Healthcare USA, Inc., New York, NY.

Hattis D, and Lynch MK. 2007b Empirically observed distributions of pharmacokinetic and pharmacodynamic variability in humans—Implications for the derivation of single point component uncertainty factors providing equivalent protection as existing RfDs (manuscript version). Available from: http://www2.clarku.edu/faculty/dhattis/ [accessed 16 April 2018]

Hill AJ, Teraoka H, Heideman W, et al. 2005 Zebrafish as a model vertebrate for investigating chemical toxicity. Toxicol Sci 86:6–19. [PubMed: 15703261]

Huangfu D, Osafune K, Maehr R, et al. 2008 Induction of pluripotent stem cells from primary human fibroblasts with only Oct4 and Sox2. Nat Biotechnol 26:1269–75. [PubMed: 18849973]

Imai K, Keele L, and Tingley D. 2010 A general approach to causal mediation analysis. Psychol Methods 15:309–34 [PubMed: 20954780]

Kameoka S, Babiarz J, Kolaja K, et al. 2014 A high-throughput screen for teratogens using human pluripotent stem cells. Toxicol Sci 137:76–90. [PubMed: 24154490]

Kang E, Wang X, Tippner-Hedges R, et al. 2016 Age-related accumulation of somatic mitochondrial DNA mutations in adult derived human iPSCs. Cell Stem Cell 18:625–36. [PubMed: 27151456]

Kavlock R, Chandler K, Houck K, et al. 2012 Update on EPA's ToxCast program: Providing high throughput decision support tools for chemical risk management. Chem Res Toxicol 25:1287–302. [PubMed: 22519603]

EPA Author Manuscript EPA Author Manuscript EPA Author Manuscript

Koturbash I, Scherhag A, Sorrentino J, et al. 2011 Epigenetic mechanisms of mouse interstrain variability in genotoxicity of the environmental toxicant 1,3-butadiene. Toxicol Sci 122:448–56. [PubMed: 21602187]

Lapasset L, Milhavet O, Prieur A, et al. 2011 Rejuvenating senescent and centenarian human cells by reprogramming through the pluripotent state. Genes Dev 25:2248–53. [PubMed: 22056670]

Lehman AJ, and Fitzhugh OG. 1954 100-fold margin of safety. Assoc Food Drug Officials US Quart Bulletin 18:33–5

Lipscomb JC, Teuschler LK, Swartout J, et al. 2003 The impact of cytochrome P450 2E1-dependent metabolic variance on a risk-relevant pharmacokinetic outcome in humans. Risk Anal 23:1221–38. [PubMed: 14641897]

Lock EF, Abdo N, Huang R, et al. 2012 Quantitative high-throughput screening for chemical toxicity in a population-based in vitro model. Toxicol Sci 126:578–88. [PubMed: 22268004]

Loomis D, and Kromhout H. 2004 Exposure variability: Concepts and applications in occupational epidemiology. Am J Ind Med 45:113–22 [PubMed: 14691975]

Maclure M 1991 The case-crossover design: A method for studying transient effects on the risk of acute events. Am J Epidemiol 133:144–53 [PubMed: 1985444]

Mann DA. 2015 Human induced pluripotent stem cell-derived hepatocytes for toxicology testing. Expert Opin Drug Metab Toxicol 11:1–5 [PubMed: 25385341]

McNally K, Cotton R, Hogg A, et al. 2014 PopGen: A virtual human population generator. Toxicology 315:70–85. [PubMed: 23876857]

Morimoto LM, White E, and Newcomb PA. 2003 Selection bias in the assessment of gene-environment interaction in case-control studies. Am J Epidemiol 158:259–63 [PubMed: 12882948]

NASEM. 2017 Controlled human inhalation-exposure studies at EPA. The National Academies Press, Washington, DC

NASEM. 2018 Environmental chemicals, the human microbiome, and health risk: A research strategy. The National Academies Press, Washington, DC

Naumova EN, Must A, and Laird NM. 2001 Tutorial in Biostatistics: Evaluating the impact of ‘critical periods’ in longitudinal studies of growth using piecewise mixed effects models. Int J Epidemiol 30:1332–41 [PubMed: 11821342]

NIH. 2018 NIH data sharing repositories. Available from: https://www.nlm.nih.gov/NIHbmic/nih_data_sharing_repositories.html [accessed 11 June 2018]

NRC. 2004 Intentional human dosing studies for EPA regulatory purposes: Scientific and ethical issues. The National Academies Press, Washington, DC.

NRC. 2009 Science and decisions: Advancing risk assessment. The National Academies Press, Washington, DC.

O’Shea SH, Schwarz J, Kosyk O, et al. 2011 In vitro screening for population variability in chemical toxicity. Toxicol Sci 119:398–407. [PubMed: 20952501]

Park SJ, Jeong SY, and Kim HJ. 2006 Y chromosome loss and other genomic alterations in hepatocellular carcinoma cell lines analyzed by CGH and CGH array. Cancer Genet Cytogenet 166:56–64 [PubMed: 16616112]

Renwick AG, and Lazarus NR. 1998 Human variability and noncancer risk assessment—An analysis of the default uncertainty factor. Regul Toxicol Pharmacol 27:3–20

Richiardi L, Bellocco R, and Zugna D. 2013 Mediation analysis in epidemiology: Methods, interpretation and bias. Int J Epidemiol 42:1511–9 [PubMed: 24019424]

Ring CL, Pearce RG, Setzer RW, et al. 2017 Identifying populations sensitive to environmental chemicals by simulating toxicokinetic variability. Environ Int 106:105–18. [PubMed: 28628784]

Rotroff DM, Wetmore BA, Dix DJ, et al. 2010 Incorporating human dosimetry and exposure into high-throughput in vitro toxicity screening. Toxicol Sci 117:348–58. [PubMed: 20639261]

Rusyn I, Gatti DM, Wiltshire T, et al. 2010 Toxicogenetics: Population-based testing of drug and chemical safety in mouse models. Pharmacogenomics 11:1127–36. [PubMed: 20704464]

Scholz G, Pohl I, Genschow E, et al. 1999 Embryotoxicity screening using embryonic stem cells in vitro: Correlation to in vivo teratogenicity. Cells Tissues Organs (Print) 165:203–11. [PubMed: 10592392]

EPA Author Manuscript EPA Author Manuscript EPA Author Manuscript

Schwartz J 2015 Identifying susceptibility in epidemiology studies: Implications for risk assessment. Available from: http://nas-sites.org/emergingscience/files/2015/10/Schwartz_Presentation.pdf [accessed 24 April 2019]

Schwartz J, Bellinger D, and Glass T. 2011 Expanding the scope of risk assessment: Methods of studying differential vulnerability and susceptibility. Am J Public Health 101:S102–S9 [PubMed: 22021313]

Schwartz J, Litonjua A, Suh H, et al. 2005 Traffic related pollution and heart rate variability in a panel of elderly subjects. Thorax 60:455–61. [PubMed: 15923244]

Scoville DK, White CC, Botta D, et al. 2015 Susceptibility to quantum dot induced lung inflammation differs widely among the Collaborative Cross founder mouse strains. Toxicol Appl Pharmacol 289:240–50. [PubMed: 26476918]

Sexton K, and Hattis D. 2007 Assessing cumulative health risks from exposure to environmental mixtures—Three fundamental questions. Environ Health Perspect 115:825–32 [PubMed: 17520074]

Shorter JR, Odet F, Aylor DL, et al. 2017 Male infertility is responsible for nearly half of the extinction observed in the mouse collaborative cross. Genetics 206:557–72. [PubMed: 28592496]

Si-Tayeb K, Noto FK, Nagaoka M, et al. 2010 Highly efficient generation of human hepatocyte-like cells from induced pluripotent stem cells. Hepatology 51:297–305. [PubMed: 19998274]

Simoni M, Baldacci S, Maio S, et al. 2015 Adverse effects of outdoor pollution in the elderly. J Thorac Dis 7:34–45. [PubMed: 25694816]

Skowronski GA, and Abdel-Rahman MS. 2001 Relevance of the 10X uncertainty factor to the risk assessment of drugs used by children and geriatrics. Hum Ecol Risk Assess 7:139–52

Studer L, Vera E, and Cornacchia D. 2015 Programming and reprogramming cellular age in the era of induced pluripotency. Cell Stem Cell 16:591–600 [PubMed: 26046759]

Swartout JC, Price PS, Dourson ML, et al. 1998 A probabilistic framework for the reference dose (probabilistic RfD). Risk Anal 18:271–82. [PubMed: 9664723]

Takahashi K, Tanabe K, Ohnuki M, et al. 2007 Induction of pluripotent stem cells from adult human fibroblasts by defined factors. Cell 131:861–72. [PubMed: 18035408]

Tan YM, Worley RR, Leonard JA, et al. 2018 Challenges associated with applying physiologically based pharmacokinetic modeling for public health decision-making. Toxicol Sci 162:341–8. [PubMed: 29385573]

Thompson CM, Sonawane B, Barton HA, et al. 2008 Approaches for applications of physiologically based pharmacokinetic models in risk assessment. J Toxicol Environ Health B Crit Rev 11:519–47. [PubMed: 18584453]

Threadgill DW, and Churchill GA. 2012 Ten years of the collaborative cross. G3 (Bethesda) 2:153–6 [PubMed: 22384393]

Threadgill DW, Miller DR, Churchill GA, et al. 2011 The Collaborative Cross: A recombinant inbred mouse population for the systems genetic era. ILAR J 52:24–31. [PubMed: 21411855]

U.S. EPA. 2001 Toxicological review of methylmercury in support of summary information on integrated risk information (IRIS). Washington, DC.

U.S. EPA. 2002 A review of the reference dose and reference concentration processes. EPA/630/P-02/002F. Washington, DC.

U.S. EPA. 2004 Toxicological review of boron and compounds in support of summary information on integrated risk information (IRIS). Washington, DC.

U.S. EPA. 2009 Integrated science assessment for particulate matter (Final report). EPA/600/R-08/139F. Washington, DC.

U.S. EPA. 2011a Toxicological review of dichloromethane in support of summary information on integrated risk information (IRIS). Washington, DC.

U.S. EPA. 2011b Toxicological review of trichloroethylene in support of summary information on integrated risk information (IRIS). Washington, DC.

U.S. EPA. 2013a Integrated science assessment for lead. EPA/600/R-10/075F. Research Triangle Park, NC.

EPA Author Manuscript

EPA Author Manuscript

EPA Author Manuscript

U.S. EPA. 2013b Integrated science assessment for ozone and related photochemical oxidants. EPA/600/R-10/076F. Research Triangle Park, NC.

U.S. EPA. 2014 Guidance for applying quantitative data to develop data-derived extrapolation factors for interspecies and intraspecies extrapolation. EPA/R-14/002, Washington, DC.

U.S. EPA. 2016a Integrated science assessment for oxides of Nitrogen—Health Criteria. EPA/600/R-15/068, Research Triangle Park, NC.

U.S. EPA. 2016b Plan to increase access to results of EPA-funded scientific research. 601-R-16-005.

U.S. EPA. 2017 Integrated science assessment (ISA) for sulfur oxides—Health Criteria (Final). EPA/600/R-17/451. Washington, DC.

U.S. EPA. 2018 Basic information about the integrated risk information system. Available from: https://www.epa.gov/iris/basic-information-about-integrated-risk-information-system [accessed 26 April 2018]

UNC Systems Genetics 2018 CC mice—Available strains. Available from: http://csbio.unc.edu/CCstatus/index.py?run=availableLines [accessed 5 June 2018]

Van de Wiele T, Vanhaecke L, Boeckaert C, et al. 2005 Human colon microbiota transform polycyclic aromatic hydrocarbons to estrogenic metabolites. Environ Health Perspect 113:6–10. [PubMed: 15626640]

Vichai V, and Kirtikara K. 2006 Sulforhodamine B colorimetric assay for cytotoxicity screening. Nat Protoc 1:1112–6 [PubMed: 17406391]

West PR, Weir AM, Smith AM, et al. 2010 Predicting human developmental toxicity of pharmaceuticals using human embryonic stem cells and metabolomics. Toxicol Appl Pharmacol 247:18–27. [PubMed: 20493898]

Wetmore BA, Allen B, Clewell HJ 3rd, et al. 2014 Incorporating population variability and susceptible subpopulations into dosimetry for high-throughput toxicity testing. Toxicol Sci 142:210–24. [PubMed: 25145659]

Wetmore BA, Wambaugh JF, Ferguson SS, et al. 2012 Integration of dosimetry, exposure, and high-throughput screening data in chemical toxicity assessment. Toxicol Sci 125:157–74. [PubMed: 21948869]

WHO. 2014 Guidance document on evaluating and expressing uncertainty in hazard characterization Harmonization project document 11. WHO, Geneva, Switzerland.

Yoo HS, Bradford BU, Kosyk O, et al. 2015a Comparative analysis of the relationship between trichloroethylene metabolism and tissue-specific toxicity among inbred mouse strains: liver effects. J Toxicol Environ Health A 78:15–31 [PubMed: 25424544]

Yoo HS, Bradford BU, Kosyk O, et al. 2015b Comparative analysis of the relationship between trichloroethylene metabolism and tissue-specific toxicity among inbred mouse strains: Kidney effects. J Toxicol Environ Health A 78:32–49 [PubMed: 25424545]

Yu J, Vodyanik MA, Smuga-Otto K, et al. 2007 Induced pluripotent stem cell lines derived from human somatic cells. Science 318:1917–20. [PubMed: 18029452]

Zeise L, Bois FY, Chiu WA, et al. 2013 Addressing human variability in next-generation human health risk assessments of environmental chemicals. Environ Health Perspect 121:23–31. [PubMed: 23086705]

Zhang J, Wilson GF, Soerens AG, et al. 2009 Functional cardiomyocytes derived from human induced pluripotent stem cells. Circ Res 104:e30–41. [PubMed: 19213953]

EPA Author Manuscript

EPA Author Manuscript

EPA Author Manuscript

Table 1.

IPCS preliminary distributions for human toxicokinetic and toxicodynamic variability.

Parameter	Median estimate across data sets of variation between 95th and 50th percentile individuals	95th percentile estimate across data sets of variation between 95th and 50th percentile individuals	Source
Toxicokinetic (TK) Variability	~2-fold	~4.5-fold	Based on variability in area under the curve from oral exposures in 37 data sets (Hattis and Lynch 2007a).
Toxicodynamic (TD) Variability	~2.5-fold	~10-fold	Based on observations of systemic, non-immune-mediated, continuous physiological parameter changes or quantal biological response in relation to internal measures of systemic exposures in 34 data sets (Hattis and Lynch 2007a).
Combined TK and TD	~3.5-fold	~14-fold	Based on Monte Carlo simulation combining the TK and TD distributions, assuming they are independent and lognormal.

Adapted from IPCS Tables 4.4 and A4.1 (WHO 2014)