Environment International 85 (2015) 120–132

Contents lists available at ScienceDirect

Environment International

journal homepage: www.elsevier.com/locate/envint

Review article

Uncertainties in human health risk assessment of environmental contaminants: A review and perspective

Zhaomin Dong, Yanju Liu, Luchun Duan, Dawit Bekele, Ravi Naidu *

The Faculty of Science and Information Technology, University of Newcastle, University Drive, Callaghan, NSW 2308, Australia
Cooperative Research Centre for Contamination Assessment and Remediation of the Environment, Mawson Lakes, SA 5095, Australia

ARTICLE INFO

Article history:
Received 6 May 2015
Received in revised form 31 August 2015
Accepted 2 September 2015
Available online xxxx

Keywords:
Uncertainties
Exposure
Hazard
Human health risk assessment
Modelling

ABSTRACT

Addressing uncertainties in human health risk assessment is a critical issue when evaluating the effects of contaminants on public health. A range of uncertainties exist through the source-to-outcome continuum, including exposure assessment, hazard and risk characterisation. While various strategies have been applied to characterising uncertainty, classical approaches largely rely on how to maximise the available resources. Expert judgement, defaults and tools for characterising quantitative uncertainty attempt to fill the gap between data and regulation requirements. The experiences of researching 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD) illustrated uncertainty sources and how to maximise available information to determine uncertainties, and thereby provide an 'adequate' protection to contaminant exposure. As regulatory requirements and recurring issues increase, the assessment of complex scenarios involving a large number of chemicals requires more sophisticated tools. Recent advances in exposure and toxicology science provide a large data set for environmental contaminants and public health. In particular, biomonitoring information, *in vitro* data streams and computational toxicology are the crucial factors in the NexGen risk assessment, as well as uncertainties minimisation. Although in this review we cannot yet predict how the exposure science and modern toxicology will develop in the long-term, current techniques from emerging science can be integrated to improve decision-making.

Contents

* Corresponding author.
E-mail address: ravi.naidu@newcastle.edu.au (R. Naidu).

http://dx.doi.org/10.1016/j.envint.2015.09.008
0160-4120/

1. Introduction

Understanding the effects of environmental contaminants on public health requires expert knowledge of the continuum from contaminant source to public health outcomes (U.S. NRC, 2009). The source-to-outcome continuum can be described as a conceptual framework for human health risk assessment (HHRA) which assimilates knowledge and techniques from chemistry, physiology, biology, mathematics, physics, medicine and other relevant disciplines. Current developments in HHRA provide important information on how to approach the release, dynamics, fate and behaviour of contaminants. However, the current paucity of knowledge concerning contaminant behaviour will inevitably result in uncertainties. Many factors including parameters, models and insufficient data, influence the frequency and degree of uncertainties (International Programme on Chemical Safety, 2005a; International Programme on Chemical Safety, 2014). The uncertainty not only points to the range and possibility of risk results to assessors, decision-makers and the public but also highlights the implications and limitations of assessment conclusions in HHRA (ENHEALTH, 2012; European Chemicals Agency, 2013; U.S. NRC, 2013).

Efforts to determine and quantify the risks posed by environmental contaminants to human health have also systemically addressed their uncertainties. In 1983, the U.S. National Research Council (NRC) published a landmark report titled 'Risk Assessment in the Federal Government: Managing the Process' (Red Book), in which the NRC not only constructed the fundamental framework for HHRA, but also documented critical uncertainties (U.S. NRC, 1983). In particular NRC recommended that those uncertainties related to the statistical, biological issues and choices of assessment and exposed population should be summarised in each step of the risk assessment process (U.S. NRC, 1983). Subsequently, the 'Guidelines for Carcinogen Risk Assessment' described the sources of uncertainties, including dose extrapolation, data and model assumption and identification for carcinogens (U.S. EPA, 1986). During the 1980s the U.S. Environmental Protection Agency (U.S. EPA) released at least five additional guidance documents to discuss how to address and determine uncertainties in HHRA for carcinogens, chemical mixtures and developmental toxicants (Williams and Paustenbach, 2002). Later, the National Academy of Sciences (NAS) emphasised uncertainty as one of several critical factors, and highlighted this issue in a series of reports published in the 1990s (U.S. NRC, 1994; U.S. NRC, 1996).

Essentially, the U.S. EPA reiterated that scientific uncertainties are unavoidable in the risk assessment process, and should be identified along with their influence on assessment. To enable management of risk from exposure to contaminants based on quantitative measures of uncertainty, U.S. EPA established a host of tools, databases and guidance protocols. For example, the benchmark dose (BMD) method, which was proposed in 1995, has been considered superior to traditional lowest-observed-adverse-effect level (LOAEL) and no-observed-adverse-effect level (NOAEL) in deriving benchmark dose values. More recently, 'Computational Toxicology Research Program (CompTox)' has been anticipated to develop a statistical toolbox assisting in reducing uncertainties (U.S. EPA, 2009). Notably, the CompTox program has developed and managed the Aggregated Computational Toxicological Resource (ACToR), a collection of databases. The ACToR aggregates data from over 500 public sources on more than 500,000 environmental chemicals, and data includes conventional toxicity, high throughput screening (HTS) and genomics tests, chemical and physical data, structural parameters, chemical identifiers and exposure database (Fowler, 2013). Meanwhile, International Programme on Chemical Safety (IPCS) has projected the harmonisation of approaches for assessing risk from exposure to chemicals: one achievement has been the application of chemical-species adjustments factors (CSAF) for interspecies differences (International Programme on Chemical Safety, 2005a; International Programme on Chemical Safety, 2005b; International Programme on Chemical Safety, 2014). In the 2014 report, 'enough knowledge' is the key criteria to judge whether the assessment is an adequate measure (International Programme on Chemical Safety, 2014). Such efforts facilitated the recognition of uncertainties in HHRAs and assisted in making decisions, and the effective utilisation of available resources potentially guarantee the appropriate 'enough' and 'coverage' we can obtain. However, due to the need to regulate many chemicals (U.S. EPA, 2014b; U.S. NRC, 2009), bridging the gap between regulation requirements and uncertainty characterisation remains a challenge.

In 2008, the next generation (NexGen) project was proposed by U.S. EPA to address advances in exposure science and toxicity testing to create an economic, rapid and more robust system for chemical risk assessment (Cote et al., 2012; Krewski et al., 2014; U.S. EPA, 2014a). *In vitro* and *in vivo* assays, molecular modelling, emerging data from genomics and proteomics all offer the possibility to understand the adverse outcome pathway (AOP), whilst novel biomarker and *in silico* simulation offers the avenue to trace chemical behaviour efficiently (Krewski et al., 2014; U.S. EPA, 2014a; U.S. EPA, 2014b). Toward a 'fit for purpose' assessments (U.S. EPA, 2014a), the aims of these new technologies not only to minimise uncertainties. Meanwhile, new technologies offer the opportunity to get to the point of making a risk decision much more efficient by rapidly characterising uncertainties. In this review, we offer a broad yet comprehensive discussion of uncertainties and solutions in the context of HHRA practice. The objectives of this review are three-fold: i) describe the uncertainties existing in HHRA processes; ii) identify methods to uncertainty characterisation; and iii) address the opportunities, challenges and approaches from emerging science that can assist in characterising and reducing uncertainties. The overview on uncertainties presented here will provide a scientific basis that contributes to a cleaner and safer environment in the future.

2. Uncertainties in HHRA

Human response to contaminant exposure depends on the toxicity of a substance and the extent of exposure. Since HHRA can usually be divided into four components, the uncertainties are also illustrated for each component as shown in Table 1.

2.1. Uncertainties in exposure assessment

Exposure assessment traces events from source to final biomonitoring, which attempts to estimate the duration, frequency, and magnitude of the exposure to a designated target group. It is anticipated that such an assessment will determine the source of toxicants, their transportation through environmental media, and their external exposure and biomonitoring under various exposure scenarios (ENHEALTH, 2012). Definitions and applications of source, external exposure and biomonitoring have been described in detail elsewhere (Zeise et al., 2013). The uncertainties in this component have been classified into scenario, model and parameter uncertainties by IPCS (International Programme on Chemical Safety, 2005b). Scenario uncertainties refer to specifying the exposure information that is consistent with the province and objective of exposure assessment, including the agent, exposed populations, spatial and temporal information, microenvironments, population activities, pathways, durations, frequencies *etc.* (International Programme on Chemical Safety, 2005b). Such elements for the scenario uncertainties are probably related to descriptive errors, aggregation errors, judgement errors and incomplete analysis (U.S. EPA, 1992).

Referring to another issue, exposure assessment techniques are usually based on an estimation of total contaminant concentration and bioavailability both of which involve uncertainties associated with methods, models and parameters. Model uncertainties stem from model structure, detail, validation, extrapolation, resolution, boundary, scenario reasonableness, etc (Jardine et al., 2003; Williams and Paustenbach, 2002). Williams et al. have summarised commonly used models currently supported and used by the U.S. EPA to assess exposures to human (Williams, 2010), and most models rely on a common

Table 1
Components, aims and sources of uncertainty in human health risk assessment.
International Programme on Chemical Safety, 2005b; International Programme on Chemical Safety, 2014; U.K. IGHRC, 2003; Williams and Paustenbach, 2002

Component	Aim	Uncertainty source
Exposure assessment	To estimate the duration, frequency and magnitude of the exposure to a target group	*Scenario*: descriptive errors, aggregation errors, judgement errors and incomplete analysis *Model*: assumptions for the correlation among exposure events; including model structure, detail, validation, extrapolation, resolution, boundary *Parameter*: in specifying the point or distribution estimate; including measurement errors, sample uncertainty, data type, extrapolation uncertainty and statistical distribution selection
Hazard identification	To determine whether the agent can cause health outcomes based on various types of evidence	Weight of multiple evidence
Dose–response assessment	To establish the dose–response curve or set the reference dose (or 'virtually safe dose')	*Database-related*: *data quality*; *heterogeneity* among studies *Extrapolation*: *extrapolating* reference dose from PoD, including interspecies, intraspecies, exposure duration, route to route; *model* selection; *distribution* assumption
Risk characterisation	To describe the nature, likelihood and magnitude from exposure to chemicals	*Decision rule*: due to toxicity criteria, site-specific scenario and parameter selection in the assessment process

Abbreviate: PoD, point of departure.

set of underlying inputs, equations, assumptions or others. Only a subset of models is able to assess uncertainties for both model input and output. Also, the sources of parameter uncertainties here include measurement errors, sample uncertainty, data type, extrapolation uncertainty and statistical distribution selection (International Programme on Chemical Safety, 2005b). In the practice of current exposure models, most input parameters are based on other existing and related investigations. For example, human exposure databases are collected from U.S. EPA's Exposure Factors handbooks, and many default scenarios and functions refer to Continuing Survey of Food Intakes by Individuals (Williams, 2010).

A sound practice to illustrate sources of uncertainties in exposure assessment is to study the reconstruction of 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD) exposure. Using research on a National Institute for Occupational Safety and Health (NIOSH) cohort as an example, the data concerning 170 workers with both estimated external exposure and serum TCDD levels (STLs) were used to derive the relationship between the two, and then this relationship was applied to reconstruct STLs over time for all 3538 workers (Steenland et al., 2001). As shown in Fig. 1, scenario uncertainties include: how to distinguish exposure from TCDD and other homologous series; how to treat the background level; and why cumulative serum liquid concentration rather than body burden has been selected as the dose metric, *etc.* On another aspect, various toxicokinetic models were adopted. While a common selection is the first-order model with a constant half-life being used for reconstructing the dose metric, evidence increasingly shows that elimination rate of dioxin is age- and dose-dependent (Aylward et al., 2005; Emond et al., 2006; Van der Molen et al., 1999). Consequently the concentration- and age-dependent model (CADM) model has been utilised for reconstructing exposure. Compared to the first-order model, TCDD body burden in CADM was distributed in liver and adipose (Aylward et al., 2005). The fraction of TCDD distribution in the two parts was correlated to TCDD concentration, and the eliminate rate was age-dependent. Differences ranging from 5- to 8-fold for reconstructed dose metric were observed between the first-order and CADM model (Aylward et al., 2005). Meanwhile, the parameter uncertainties involve the determination of half-life for the target group, exposure matrix, background level, *etc.*

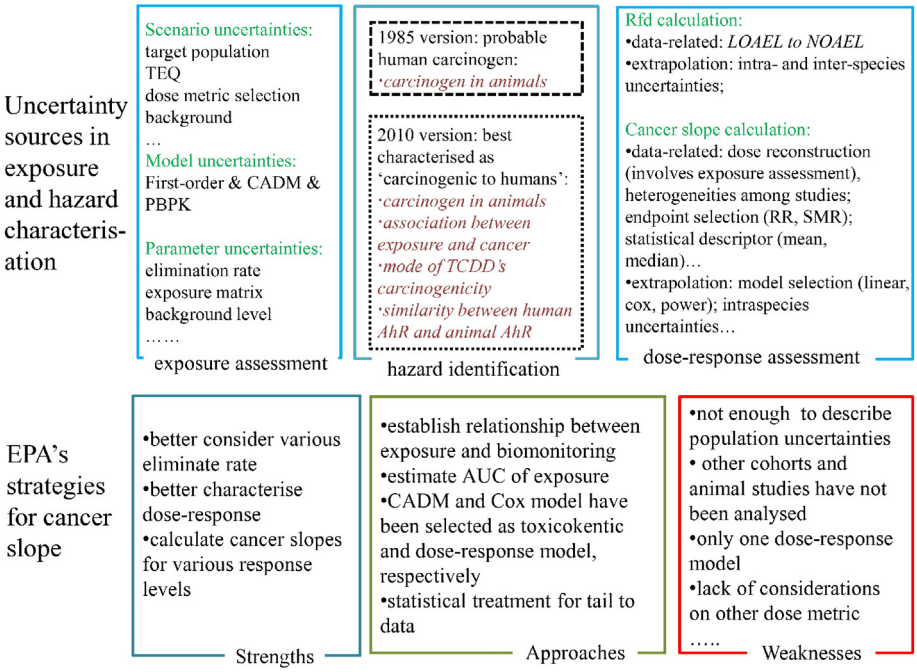

Fig. 1. Uncertainty sources in exposure assessment and hazard characterisation, and EPA's strategies for cancer slope: case study on TCDD (Cheng et al., 2006; U.S. EPA, 2010). Abbreviation. TEQ: toxic equivalency factor; CADM: concentration- and age-dependent elimination model; PBPK: Physiologically-based Pharmacokinetic; RR: relative risk; SMR: standardised mortality ratio; LOAEL: lowest observed averse effect level; NOAEL: no observed adverse effect level; AUC: area under curve.

2.2. Uncertainties in hazard identification

Hazard identification attempts to determine whether the agent is able to cause adverse health outcomes based on various types of evidence for toxicity. In 1986, based on the scientific evidence from animal, human epidemiological, and mechanism or mode-of-action studies, U.S. EPA classified cancer chemicals into five categories: human carcinogen, probable human carcinogen, possible human carcinogen, not classifiable as to human carcinogenicity, and evidence of non-carcinogenicity to humans (U.S. EPA, 1986). Key issues involved in this procedure include how to weigh positive and negative results, crucial research, and sensitivity endpoints for final judgments. In the 1986 cancer guideline, the weight of evidence was tied to specific types of data (*i.e.* epidemiology, animal, *etc.*) (U.S. EPA, 1986). In contrast to the step-wise approach in the 1986 version, current cancer guidelines emphasise the importance of weighing all of the evidence in reaching conclusions about the human carcinogenic potential of agents (U.S. EPA, 2005). A common issue with uncertainties is the debate on their nature, reliability, relevance and consistency with reference to animals and humans.

Except uncertainties when identifying hazard from human and animal studies, uncertainties exist when using the ancillary information from *in vitro* assays or quantitative structure–activity relationships (QSAR) to help with identification. *In vitro* tests are not reliable surrogates for a complete biological system, and more importantly, a finding from an *in vitro* test is not totally predictive for *in vivo* results. The QSAR model which is a regression or classification method used to establish relationships between chemical structure (or properties) and biological activity (or quantitative biological) that in turn predicts the toxicity of studied compounds, was proposed one hundred years ago (Schultz et al., 2003). Chemicals can be grouped according to their structural features or the fragments that may be responsible for the toxicity endpoint. The uncertainty here limited the prediction and application of the QSAR model when support from mechanistically relevant biological data is lacking (Rusyn and Daston, 2010).

The classification of TCDD well illustrates how to consider the weight of available evidence in hazard identification. In 1985, based on long-term bioassays conducted in both sexes of several strains of rats and mice, hamsters, and fish, U.S. EPA demonstrated TCDD is a carcinogen in laboratory animals. However, at that time, the epidemiological evidence for the carcinogenicity of TCDD alone was inadequate, and no epidemiological evaluation was available (U.S. EPA, 1985). Thus, in 1985, U.S. EPA classified TCDD as a "probable human carcinogen (B1)" based on the overall weight of evidence (U.S. EPA, 1985). Later, data concerning several epidemiological studies, especially information from the NIOSH cohort, Hamburg cohort and BASF cohort, was reported to support the classification for 2,3,7,8-TCDD. Although the data are insufficient to establish a causal link to dioxin exposure (Starr, 2001), these occupational studies have strongly and consistently demonstrated the association between 2,3,7,8-TCDD exposure and cancer risk (Crump et al., 2003). Together with the carcinogenicity evidence from a number of species, mode of TCDD's carcinogenicity (Aryl hydrocarbon receptor (AhR)-promoter), the similarity between human AhR and animal AhR, U.S. EPA concluded that TCDD was "best characterized as 'carcinogenic to humans'" in 2003 (U.S. EPA, 2003).

2.3. Uncertainties in dose–response assessment

The aim of dose–response assessment is to establish the dose–response curve or extrapolate the reference dose (Rfd) (non-cancer effect). The Rfd is an estimate of a daily oral exposure based on an acceptable risk during a lifetime. Similarly, this concept for a continuous inhalation exposure to the human population is defined as reference concentration (Rfc). Especially, the key of this process is to ensure enough coverage of Rfd for sensitive populations, such as infants, children, females, elders, and pregnant groups. Regarding the cancer effect, cancer slope factors (CSF) – usually a 95% upper confidence limit – are used to estimate the risk of cancer associated with exposure to a carcinogenic or potentially carcinogenic substance (U.S. EPA, 2014b). The Rfd and CSF provide the basis for environmental management, such as risk estimation, chemical priority or guidance setting. Across the span of uncertainties, the toxicity value may vary across an order of magnitude. Based on uncertainties, the confident interval on incidences of risk can convey explicit information on chemicals' toxicity. For example, U.S. EPA evaluated that the risk of cancer from drinking water containing arsenic is 5E − 5 per (μg/L). Considering uncertainties, the skin cancer risk is estimated to range between 1E − 3–2E − 3 for a lifetime's arsenic exposure to drinking water (concentration is 1 μg/kg/day, consumption is 2 L for a 70 kg person) (U.S. EPA, 1988).

Typically, two types of uncertainties are involved: determination of point of departure (PoD) and then extrapolating PoD to human Rfd (or CSF) across the species, population, routes, doses and durations. And such two types of uncertainties can be summarised as database and related extrapolation (Table 1) (U.K. IGHRC, 2003).

Two approaches are usually utilised for determining PoD: the NOAEL/LOAEL and BMD approaches. The NOAEL/LOAEL approach is experimentally derived and it depends on the number of doses and 'tightness' of the range of doses which are included in a given experimental design. Many criticisms have been made of its limitations: firstly, NOAEL is based on judgement; secondly, experiments involving fewer animals tend to produce larger NOAELs; thirdly, there is difficulty in determining slope; and fourthly, failure to estimate size of the 'population at risk' when connecting with exposure data (U.S. EPA, 1995). Compared to NOAEL/LOAEL, the BMD approach incorporates and conveys information from all the designated experimental groups, and then calculates the BMD based on the statistical theory. A two-step process is used to assess the dose–response model for estimating the BMD: 1) obtaining the observed data based on various sources, including animal studies, *in vitro* assay study, and math modelling; and 2) establishing the dose–response curve within the range of observations. Likelihood method has been employed to construct the dose–response curve in order to evaluate BMD, as well as lower confidence limit for BMD (BMDL). In the latest version of BMD software, about 30 different models are available for the analysis of dichotomous data, continuous data, nested developmental toxicology data, multiple tumour analysis and concentration–time data. Using a hybrid approach, the generic BMD can be applied to some non-standardised data, including the case for quantal end-points.

During the extrapolating from PoD to equipotent human dose, interspecies and intraspecies uncertainties are probably the most important factors. Species differences stem from metabolic, functional and structural variations (International Programme on Chemical Safety, 2005a). Specifically, the interspecies can be divided into two aspects: toxicokinetics and toxicodynamic differences. While interspecies describe the differences of dose–response between animals and human, the intraspecies uncertainties account for variations in sensitivity within the human population (International Programme on Chemical Safety, 2014). Both the interspecies and intraspecies differences can be divided into toxicokinetics (TK) and toxicodynamics (TD), which allow each component to be determined by relevant, available chemical-specific data.

In addition, uncertainties do exist when extrapolating high dose in animal experiments to low exposure for humans, as well as the extrapolations between different exposure routes, durations (such as acute exposure to chronic exposure) and modes. Too often, there are insufficient low doses included in dose–response relationship studies, as well as exposure duration. Bokkers has demonstrated a ratio of 2–5 in the chronic BMD when using a shorter-duration BMD (Bokkers and Slob, 2005). Meanwhile, the current BMD version uses Akaike Information Criterion (AIC) to select or compare models. However, a mathematical choice may not account for the biological mechanism, and using a specialist judgement or other strategy as the prior information may improve the choice (An et al., 2009; Johnson and Omland, 2004).

Still using the TCDD as a case study (Fig. 1), U.S. EPA has considered both the calculation for Rfd and cancer slope. In 2012, U.S. EPA established the Rfd based on two epidemiological studies that associated TCDD exposures with: a) decreased sperm concentration and sperm motility; and b) increased thyroid-stimulating hormone levels. The LOAEL has been estimated to be 20 pg/kg/day, and two types of uncertainties have been considered here to final calculation of Rfd: firstly, extrapolation from LOAEL to NOAEL (determination of PoD); and secondly, intraspecies differences. With reference to estimating the cancer slope, U.S. EPA has employed the upper 95% bound on the regression slope and assumed that this confidence interval is enough coverage for the sensitivity group (U.S. EPA, 2010).

2.4. Uncertainties in risk characterisation

Risk characterisation is a process that describes the nature, likelihood and magnitude of adverse effects by integrating toxicological and exposure information. Selections of data and models result in multiple options on the estimated risk, and a sophisticated HHRA should retain all materials employed in the entire process for further censoring (ENHEALTH, 2012). Uncertainties here, defined as decision-rule uncertainties, would be propagated due to the exposure scenario and toxicity selection. A case is the World Health Organization (WHO) utilising the disability-adjusted-life-year since the 1990s to assess the burden of disease consistently across diseases, risk factors and regions. During this process, the choices of the disability weights, discounting and age weighting affected the reliability of the outcomes, as demonstrated by a several-fold division of the risk that was observed when the discounting was assumed to be 3% and zero for life-time exposure. The quantitative and qualitative uncertainty characterisation in risk characterisation would firstly, inform the risk manager about the confidence in the data, and secondly, describe benefits to the greatest extent possible.

3. Classical methods for uncertainty characterisation

Compared to variability, a common feature of uncertainty is that while it can be identified it is difficult to quantify due to data limitation and mechanism deficiency. Currently, the tiered approach has been suggested for uncertainty characterisation according to the available information. For example in exposure assessment, such uncertainty characterisation can be divided into four tiers (International Programme on Chemical Safety, 2005b): a) tier 0, screening level to determine the exposure or risk is acceptable based on exposure assumptions; b) tier 1, qualitative level to address the key sources contributing to exposure assessment; c) tier 2, deterministic level to estimate the risk range and contributions from involved parameters; and d) tier 3, probabilistic level characterising the individual and combined effects of input and parameter uncertainties on the outcomes. For a given scenario, the first step is to ascertain which level of uncertainty characterisation can be obtained based on available knowledge. Regardless of which level is selected, it is important to convey information on distinguishing reducible from irreducible uncertainty, separating individual variability and true scientific uncertainty to the decision-maker, and further to minimise uncertainty or build confidence for assessment results (U.S. NRC, 2009).

In 1983, NRC recommended inference guidelines or defaults for the treatment of uncertainty, which was followed in 1994 by NRC providing recommendations of default and quantitative methods for uncertainty propagation (U.S. NRC, 1983; U.S. NRC, 1994). The 'use of default' includes the model selection, distribution assumption and parameter choice. A default option appears to be the best choice in the absence of data to the contrary based on a relatively strong scientific foundation (U.S. NRC, 2009). A high level of uncertainty analysis – such as quantitative uncertainty analysis – requires additional technique and data, including advanced statistical analysis in the form of the Physiologically-based Pharmacokinetic (PBPK) model. Several tools used for both quantitative and qualitative uncertainty characterisation, as well as data requirements, are summarised in Table 2. In practice, these tools lie on a continuum and combination that can be modified in a generic approach for reducing uncertainties across different components in HHRA, such as the integration of Bayesian technique and PBPK model to update our knowledge of tetrachloroethylene chloroethylene (Gelman et al., 1996).

3.1. Uncertainty factors

Table 3 summarises the default uncertainty factors (UF) used in hazard characterisation from determination of PoD to the establishment of Rfd. It also outlines some practices in the Integrated Risk Information System (IRIS). Specifically, the total number of uncertainty factors according to data quality results can range from 10 to 3000. Using TCDD as a case study, while U.S. EPA assumed a total UF to be 30, most other organisations or countries (including IPCS, the Joint FAO/WHO Expert Committee on Food Additives, Japan, United Kingdom) have selected value 10 as the UF (Junko and Isamu, 2008). In Japan, the UF for extrapolating from LOAEL to NOAEL has been assumed to be 3, while U.S. EPA has defaulted this UF to be 10. The data source includes human studies, animal studies and mechanistic studies, and the common strengths and weaknesses of these three types have been summarised by NRC (U.S. EPA, 2014b).

An uncertainty factor of 10 or 100 is typically assumed for the interspecies and intraspecies differences as shown in Table 3 (Lehman, 1954). However, this factor is just a safety margin without a scientific or mathematical basis. In 1966, Freireich et al. suggested that species'

Table 2
Classical methods and their applications for characterising uncertainties in human health risk assessment.

Tool	Application	Data required
Guidance	Provide standard operation system for compliance scheme; reduce uncertainties from measurement	Low
Expert judgement	Offer experienced knowledge and consensus for rule decision	Low
Uncertainty factor	Use the empirical factors for extrapolation	Low
CSAF	Use of data in dose–response assessment for interspecies and intraspecies uncertainties	Medium
Biomonitoring	Use the internal data to indicate risk and understand exposure scenario	Medium
Benchmark dose	Reduce uncertainties by using utilisation of NOAEL and LOAEL	High
BBDR model	Link external exposure to an adverse effect by incorporating data on biological processes at the cellular and molecular level	High
PBPK model	Simulate tissue concentration under various exposure scenarios; reduce uncertainties from inter-species and intra-species	High
Monte Carlo Simulation	Describe population variability and uncertainties	Medium
Sensitivity analysis	Determine sensitivity parameters crucial to variability and uncertainties of risk results	Medium
Model averaging	Weigh the fitness of various models; reduce uncertainties from model selection	Medium
Meta-analysis	Distinguish between real effects and heterogeneities among studies	High
Bayesian technique	Optimise input parameters by incorporating prior information; reduce uncertainties from model parameters	High

Abbreviations: PBPK: Physiologically-based Pharmacokinetic; BBDR: biologically based dose–response; CSAF: chemical-specific adjustment factors; NOAEL: no observed adverse effect level; LOAEL: lowest observable adverse effect level.

Table 3
Summary of applied uncertainty factors for selected chemicals.

	UF1 Determination of PoD	UF2 Interspecies	UF3 Intraspecies	UF4 Exposure duration	UF5 Exposure pathway	UF6 Others
General description	1. NOAEL/LOAEL 2. BMD	Extrapolation across species	Enough coverage to protect sensitivity group	Extrapolation across subchronic, subacute, chronic exposure	Extrapolation across oral, inhalation, skin or others	Data quality, model selection, distribution assumptions
Default values	BMD/BMDL: 1 LOAEL/NOAEL: 3–10	Default: 10 Body size$^{3/4}$: 1–10	Default: 10	Subchronic–chronic: 2 Subacute–chronic: 5	1/RDDR or 1/RGDR	Case-specific
Methylmercury (Integrated Risk Information System, 2001)	PoD: $BMDL_{0.5}$ UF1: 1	1	Toxicokinetic: 3.16 toxicodynamics: 3.16	/	/	/
Data based on human study, PoD is 0.857–1.472 μg/kg bw per day, total UF is 10, and thus Rfd is established to be 0.1 μg/kg bw per day						
2,3,7,8-TCDD (U.S. EPA, 2012a)	PoD: LOAEL UF1: 10	1	3	/	/	/
Data based on human study, PoD is 20 ng/kg per day, total UF is 30, and thus Rfd is established to be 0.7 pg/kg bw per day						
Trichloroethylene (Keil et al., 2009)	PoD: LOAEL UF1: 10	Toxicokinetic: 1 toxicodynamics: 3.16	Toxicokinetic: 1 toxicodynamics: 3.16	/	/	/
Data based on female B6C3F1 mice, PoD is 48 μg/kg bw per day, total UF is 100, and thus Rfd is established to be 0.48 μg/kg bw per day						
Deoxynivalenol (International Programme on Chemical Safety, 2014)	PoD: LOAEL UF1: 3	10	10	/	/	/
Data based on B6C3F1 mice, PoD is 120 μg/kg bw per day, total UF is 300, and thus Rfd is established to be 0.4 μg/kg bw per day						
Bisphenol A (National Toxicology Program, 1982)	PoD: LOAEL UF1: /	10	10	Subchronic to chronic: 10	/	/
Data based on rat study, PoD is 50 mg/kg per day, total UF is 1000, and thus Rfd is established to be 0.05 mg/kg bw per day						
2,2′,4,4′,5-Pentabromodiphenyl ether (Viberg et al., 2004)	PoD: BMD_{1SD} UF1: 1	10	10	Single-dose to lifetime exposure: 3	/	Database deficiency: 10
Data based on C57/B1 mice, PoD is 0.29 mg/kg per day, total UF is 3000, and thus Rfd is established to be 0.1 μg/kg bw per day						

Abbreviations: BMD_{1SD}: maximum likelihood estimate of the dose corresponding to a change in the mean equal to one SD of the control mean; PoD: point of departure; Rfd: reference dose; NOAEL: no observed adverse effect level; LOAEL: lowest observable adverse effect level; RDDR: regional deposited dose ratio; RGDR: regional gas dose ratio; UF: uncertainty factor.

different toxicity responses depend on the external surface area of the respective species rather than on their weight (Freireich et al., 1966). Hence, the U.S. EPA recommended using the three-quarter power of the weight ($w^{3/4}$) as the uncertainty factor to adjust the reference dose (U.S. EPA, 1986; U.S. EPA, 1996) and consequently a factor of 1–10 has always been applied for extrapolating the equipotent human dose from rat or mouse studies. As knowledge increases, appropriate chemical-specific data and information would modify this default approach, with the best strategy being to use the CSAF (U.S. EPA, 2011). IPCS has concluded the TK to be 4.0 and TD to be 2.5, since it was based on underlying physiological differences between species (International Programme on Chemical Safety, 2005a; Renwick, 1993). The estimate for the clearance based on *in vivo* studies can help determine the TK ratio across the species. Regarding the intraspecies differences, the TK and TD values are assumed to be 3.16 ($10^{0.5}$) and 3.16 ($10^{0.5}$), respectively (International Programme on Chemical Safety, 1999). This subdivision was supported by the analysis of a database on the variability in both body (kinetics) and target organ sensitivity (dynamics) (Renwick and Lazarus, 1998). It is therefore possible to modify the values based on available information. For example, recent reviews using drug databases concluded that the current general kinetic value (3.16) cannot account for intra-species uncertainties (such as CYP2D6) among the sub-population groups (Dorne and Renwick, 2005; Dorne et al., 2005).

As stated above, use of $BW^{3/4}$ scaling in combination with a reduced default interspecies uncertainty factor is recommended as the default approach to replace the previous default approach with a full uncertainty factor (*i.e.* a UF value of 10) (U.S. EPA, 2011). For the purpose of creating a hierarchy of approaches to derive Rfd, $BW^{3/4}$ scaling method may not always predict oral exposures associated with precise toxicologically-equivalent doses for specific chemicals. An intermediate approach which uses the CSAF to help reduce UF was endorsed by U.S. EPA (U.S. EPA, 2011). Case studies employed applicable data (interspecies differences and human variability in toxicokinetics and toxicodynamics) to illustrate the principles of development of CSAF (International Programme on Chemical Safety, 2005a). Recently, to incorporate variability in the parameter, the HD^{I}_{M} (the human dose where a fraction I of the population shows an effect of magnitude) approach was recommended recently in the 2014 IPCS harmonisation project report (International Programme on Chemical Safety, 2014). While this novel approach still employed uncertainty factors to extrapolate Rfd, it is still powerful to inform population variability.

3.2. Physiologically-based Pharmacokinetic (PBPK) model

Integrating toxicokinetic data or data into the mode or mechanism of action in the dose–response or concentration–response model will help supersede the default value. The PBPK model, describing absorption, distribution, metabolism and excretion (ADME) of synthetic or natural chemical substances in biological systems, can be used for dose estimation, supporting extrapolation, exposure reconstruction, experimental design, and hypothesis and theory testing. As illustrated in Fig. 2, a PBPK model envisages the body as a set of similar compartments that represent the organs or tissue groups. Implications for these physiological compartments are the important differences emerging between the classical compartment and PBPK models, whereby the latter can simulate biomarker concentrations in various species scenarios. One of the most widespread applications of PBPK models is to forecast the impact of specific mechanistic processes and determinants on the tissue dose. In fact the *in vitro* data and *in silico* simulation can be integrated for making predictions of the tissue dosimetry in the whole animal, and further extrapolating this to humans (Lipscomb et al., 2012).

Developments in computational science have boosted the application of PBPK model in HHRA. Clark et al. reviewed the principal

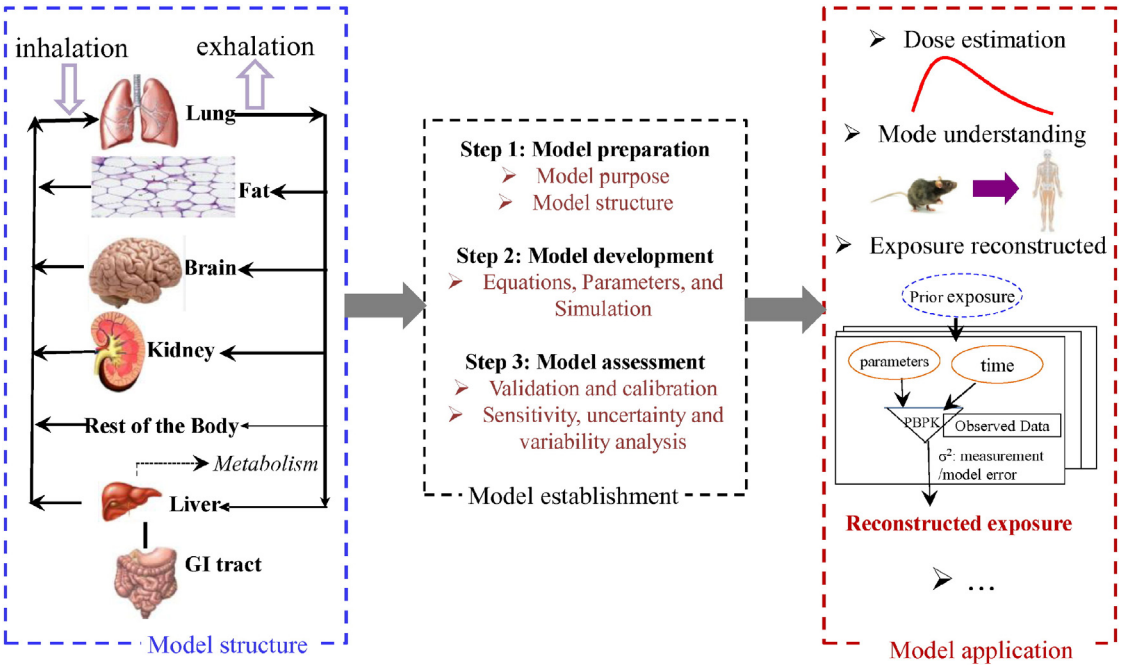

Fig. 2. PBPK model structure, process for PBPK model establishment and PBPK model application. Abbreviation. PBPK: Physiologically-based Pharmacokinetic.

processes for documenting a PBPK model (Clark et al., 2004), and here we have re-organised these into three steps (Fig. 2). Firstly, model preparation includes identification of the model's structure and purpose, including data interpretation, pharmacokinetic parameters estimate, and dosimetry prediction, *etc.* Secondly, model development is where the mathematical equations, model parameters and computer code are developed. Most model equations are developed based on mass balance. Generally, parameters used in a PBPK model can be classified as physiological parameters (*e.g.* body weight, tissue blood flow rate, and cardiac output), partition coefficients (the ratio of chemical concentration in tissue and blood) and biochemical parameters (metabolism, absorption, and excretion). Commercial software applications such as acslX, Matlab and Berkeley Madonna have been widely applied to model programming. Thirdly, model assessment includes model evaluation and calibration, sensitivity analysis, uncertainty and variability analysis. This process is required to assess the suitability and applicability of a PBPK model, which will improve the model's reliability and application.

A mechanistic understanding of the PBPK model can potentially reduce uncertainties arising out of hazard characterisation and exposure assessment by specifying parameters across species, pathways and chemicals. When utilising the PBPK model, a resulting human-equivalent dose (7.6 mg/kg/day) was determined for ethylene glycol monobutyl ether associated with the same level of the dose metric in animal (Lipscomb et al., 2012). Based on the collection of physiological, chemical, *in vitro*, and *in vivo* data, Chiu et al. developed a comprehensive, Bayesian, PBPK model-based analysis of the population toxicokinetics of trichloroethylene (TCE) and its metabolites across species (Chiu et al., 2009; Chiu et al., 2006). The established PBPK model has confirmed that mice had the greatest rate of respiratory tract oxidative metabolism compared to rats and humans (Chiu and Ginsberg, 2011; Chiu et al., 2009). Since using the PBPK models accounts for toxicokinetic differences between animals and humans, as well as normal groups and sensitivity groups, the interspecies and intraspecies has been reduced from a default value of 10 to a modified one, *i.e.* 3 (Keil et al., 2009). The main purpose of the PBPK model should be to continuously focus on how to identify possible featured structures and determine substance-specific parameters, particularly metabolism parameters and 'population' parameters.

3.3. Statistical tools

Characterisation of probability and shape of risk are both necessary, since the probability of risk can offer a point estimate of the expectation risk while the shape of risk can outline and compare risks for subgroups, especially for the vulnerable group. The Organisation for Economic Co-operation and Development (OECD) has summarised most statistical issues associated with the design and analysis of chronic toxicity and carcinogenicity bioassays, the analysis and interpretation of tumour data and their use (Organisation for Economic Co-operation and Development, 2010). It is noted that the treatment of data analysis should not be tied to one single approach, and it is advisable to propose new and modified approaches, for example artificial neural networks and general linear analysis.

3.3.1. Monte Carlo Simulation (MCS)

. Of all the available techniques, MCS is the most basic one used in HHRA and it has been widely applied in exposure assessment, population PBPK model and Bayesian simulation (Gelman et al., 1996; Glorennec et al., 2007; Yang et al., 2010). Compared to point estimate, interval estimation based on MCS can improve our understanding of contaminants' environmental behaviour and address existing uncertainties. By assuming a distribution (usually normal, log-normal, uniform, gamma distribution) for the involved parameters, results can be reported with interval or boundary as a frequency graph. Theoretically, selecting candidate probability distributions should be done according to their underlying physical processes or mechanisms. For example log-normal distribution is usually assumed for biomarker concentrations, and gamma distribution for random parameters of interest is the sum of independent exponential parameters. For correlated parameters, multivariate distributions are usually employed to include parameter correlations (Axelrad et al., 2007). In 1997, U.S. EPA established a general framework and broad set of principles for MCS practice quantitative risk assessment (U.S. EPA, 1997). It is worth noting that a relatively large dataset is required. Risk assessors should confirm that there is enough data to make the necessary decision before implementing MCS.

In some situations, MCS does not treat variability and uncertainty separately, and thus the assuming distributions will reflect both the

inherent parameter heterogeneity and uncertainty about the accuracy of measurements (Poulter, 1998). It will therefore fail to distinguish between variability and uncertainty. Meanwhile, since vulnerable groups are of the greatest interest, the tails of risk distributions are sensitive to the shape of input distribution. In estimating the low-dose risk, the differences are observed between the gamma model and log-normal model, and their possible range over several orders of magnitude has been estimated (Crump et al., 2010b). Nevertheless, the application of MCS to propagate uncertainty for input variables to output is relatively straightforward and may be valuable to the consumer of the information, particularly when it is extended using sensitivity analysis (Poulter, 1998).

3.3.2. Sensitivity analysis

Sensitivity analysis is the process of providing a ranking of input parameters based on their relative contributions to output variability and uncertainty. Generally, sensitivity analysis provides a quantitative examination to determine sensitive variables crucial to the outcome of a risk assessment, which further aids in probabilistic methods to represent uncertainties. In the U.S. EPA's risk guidance procedures, sensitivity analysis is considered to be a basic and necessary component for risk characterisation as exemplified in both cancer and non-cancer practice (U.S. EPA, 2000; U.S. EPA, 2001b; U.S. EPA, 2002; U.S. EPA, 2005). Many analytical approaches are referred to as forms of sensitivity analysis, and are usually generalised as either tier 1 or tier 2 approaches (U.S. EPA, 2001a). The tier 1 approach for sensitivity analysis relies on point estimate, by observing output when only changing the target parameter. However, it fails to address interrelationships among underlying parameters when input variables are not independent. In tier 2 when employing MCS, sensitivity analysis enables input variables to vary simultaneously. Further sensitivity results based on regression analysis overcome the limitations from the strategy used in tier 1.

3.3.3. Meta-analysis

Glass defined meta-analysis as "the statistical analysis of a large collection of analysis resulting from individual studies for the purpose of integrating the findings" (Glass, 1976). Meta-analysis is therefore a statistical procedure for amalgamating, summarising and reviewing previous quantitative data that offers a more precise estimation of a treatment outcome than what a single source of quantitative data has estimated. It may explain discrepancies that exist in various individual studies' results (Egger et al., 1997). Specifically, meta-analysis integrates the results of many studies and subsequently increases the available data and confidence to test the null hypothesis. In HHRA an upsurge occurred in the development and application of meta-analysis during the 1990s due to an expanding need for reliable summaries of health and medical research. In the practice of uncertainty minimisation, a meta-analysis distinguishes the real effects of heterogeneity among a cluster of studies.

Over the past twenty years great strides have been made in the development and refinement of statistical approaches in the context of meta-analysis, and the issues involving random effects, Bayesian inference, hypothesis analysis, trend analysis, *etc.* Cases studies on mercury, lead, cadmium, PAHs and other contaminants using this technique highlight gaps in our hazard description knowledge (Armstrong et al., 2004; Axelrad et al., 2007; Gallagher and Meliker, 2010; Lanphear et al., 2005). A major limitation is data availability. Most data used in meta-studies are selected from literature or reports and the probability of a study being published mostly depends on the statistical significance of its results (Whitehead, 2003), which is termed publication bias. Various methods including Egger's method, truncated sampling and weighted distribution theory, have been proposed and applied to correct publication bias; however, they all have their limitations (Thornton and Lee, 2000). Obtaining data from all relevant randomised studies is ideal but difficult to achieve.

3.3.4. Bayesian Hierarchical Model (BHM)

Recent developments in Bayesian inference, as an extension of MCS, afford an alternative approach for quantifying uncertainties and reducing the default factors used (U.S. EPA, 2006). In accordance with Bayesian theory, reconstructed information of an objective parameter is obtained from the product of joint prior information and a likelihood function, whose form is based on the measurement model. The core of the BHM model is using the likelihood function which is built on the differences between monitoring data and simulated results and prior information to obtain posterior information. Combining with the PBPK model, the application of Bayesian analysis to update population parameters has been rapidly developed in HHRA (Gelman et al., 1996; Yang et al., 2010). Five key features were proposed during the procedure: 1) a physiological mode for linking exposure and biomarker; 2) a population model for accounting for population variability and uncertainties; 3) prior information for the objective parameters; 4) observed data for developing the error model between the simulation result and experimental data; and 5) Bayesian inference (Gelman et al., 1996). Compared to the deterministic approach, at least three properties can be addressed by Bayesian inference: distribution, convergence and inter-correlation (Xu et al., 2006). In additional, Bayesian updating probably helps the mechanism for understanding chemical behaviour. Based on a research study using BHM of BDE-209 in Chinese sturgeon, results showed distribution rather than the absorption or metabolism, is a major factor affecting the bioaccumulation of decabrominated diphenyl ether (Wan et al., 2013). Since it usually involves a heavy workload of population parameter estimations and model calculations in many dimensions, the data requirement is much greater than when employing traditional approaches.

In the exposure reconstruction for the NIOSH cohort (see Fig. 1), the CADM model rather than the first-order model has been selected (Aylward et al., 2005). In the meantime a Cox model has replaced previous versions (Cheng et al., 2006). Treatments with fewer uncertainties were authorised by U.S. EPA, and an oral cancer slope factor of 10^6 per (mg/kg-day) has been established when the target risk range is 10^{-5} to 10^{-7} (U.S. EPA, 2010; U.S. EPA, 2012a). Although the developed PBPK model for TCDD has not yet been applied due to the lack of PBPK parameters in this case (Emond et al., 2006), a series of reports from U.S. EPA on 'platform chemicals' like TCDD attempted to integrate continuous and key scientific evidence that addresses uncertainties. Yet it provides what these chemicals consider to be an adequate margin of safety. Using the illustration from the TCDD case study, it should be recognised that the principle goal of HHRA is not to pursue 'accuracy', but rather to achieve an adequate measure by avoiding the understatement of risk to protect public health. Central to the classical approaches is the fact that they are based on available data streams that minimise reducible uncertainties, as well as outline irreducible fraction. On the other hand, current strategic decisions are still largely limited by the paucity of data, varying recurring issues in HHRA, new contaminants, *etc.*

4. Perspectives on the emerging science

The regulatory requirement is still being developed, which is beyond the spectrum of traditional approaches and thus contributes to uncertainties in the HHRA. Emerging science has the potential to address these challenges and uncertainties by obtaining and explaining available data streams. For example, beginning on 1st June 2007, the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH) attempted to estimate the circulation of substances on the market. REACH considers chemicals that have high yield, carcinogens, mutagens or substances toxic to reproduction (CMB), persistent bioaccumulative and toxic substances (PBT) and very persistent and very bioaccumulative substances (vPvB) as priority chemicals. In 2012, U.S. EPA updated information on the commercial manufacturing, processing, and use of about 7690 chemicals in the United States (U.S.

EPA, 2012b). To assess the potential risk posed by huge emission inventories, it is anticipated that modern exposure science will access exposure databases (including emission inventories, environmental databases and demographic and activity databases). It must also ensure that rapid, efficient and predictive capabilities are invested in: firstly, a toolbox for screening, prioritising and prediction; and secondly, the actual tools developed in this area (Lioy, 2008; Lioy, 2010; Lioy and Smith, 2013).

Additionally, traditional descriptive and mechanistic toxicology is too labour- and resource-intensive to comprehensively evaluate contaminants (Rusyn and Daston, 2010). In response to these challenges, U.S. NRC released a report titled 'Toxicity Testing in the 21st Century in 2007' (U.S. NRC, 2007). In this report, four criteria were outlined for a paradigm for toxicity testing: 1) broad coverage of chemicals; 2) timeliness and cost-effectiveness; 3) a robust scientific basis; and 4) reduced testing of animals (Andersen and Krewski, 2009).

In 2011, NexGen embodied a three-tier framework for risk assessment from screening, ranking and prioritising thousands of chemicals with little data other than high-throughput, *in vitro* testing, QSAR and read-across to augmenting traditional, data-rich chemical assessments for regulatory decision-making (Cote et al., 2012; Krewski et al., 2014; U.S. EPA, 2014a). Based on this framework, NexGen attempts to address the pivotal issues in HHRA (U.S. EPA, 2014a): individual and population-level effects (International Programme on Chemical Safety, 2014), human variability and susceptibility (Zeise et al., 2013), mixtures and non-chemical stressors (Callahan and Sexton, 2007), interspecies extrapolation (Burgess-Herbert and Euling, 2013), and low-exposure level exposure (Rhomberg et al., 2011). To achieve the aims in the respective tiers, tools and technique requirements and understanding of the mechanism increase when moving from tier 1 to tier 3, and uncertainties decrease.

To generate, interpret and address more available data, a range of tools including high-throughput screens (HTS), stem cell biology, function genomics, bioinformatics, system biology, computational system biology, PBPK models, QSAR models, and biomarkers, have been applied and articulated in a previous study (Andersen and Krewski, 2009). However, of all the tools, PBPK has been the most widely applied in quantitative risk assessment, while QSAR and HTS have been widely employed for prioritising chemicals (Organisation for Economic Co-operation and Development, 2009; Wambaugh et al., 2013). For example, it has been proposed that quantitative HTS (qHTS) can assist in chemical screening and prioritisation. qHTS is able to obtain a dose–response curve across a broad concentration range. In 2006, the plan was to evaluate the activities of 2870 compounds using 1536-well plate qHTS (Tice et al., 2013). Tier 2 is anticipated to be based on meta-analysis of bioinformatically identified and organised existing data (both *in vivo* and *in vitro*, and short-duration exposure *in vivo* bioassays data), while tier 3 will employ robust traditional data augmented with new data types for decision-making (Cote et al., 2012). In 2014, U.S. EPA developed eight prototype assessments for illustrating how the cited new data and approaches can be applied to HHRA, as well as their strengths and weaknesses (U.S. EPA, 2014a). Following the tiered approach in the NexGen HHRA, the comparisons of key uncertainty characterisation in the current practice and NexGen paradigm are summarised in Table 4. The essence of emerging science is to provide rich data streams and computational tools for uncertainty characterisation. Table 4 highlights the potential applications of biomonitoring, *in vitro* data and computational toxicology in filling the gap concerning uncertainty reduction in the NexGen HHRA.

4.1. Biomonitoring

There have been significant advances in techniques to obtain available data. Despite this development, however, exposure information for 95–99% of the 100,000 chemicals having toxicity information, is still unavailable (Egeghy et al., 2012). Traditional exposure assessment focuses on using measurement, analysis and models to estimate the degree and pathways of exposure. Compared to exposure assessment, the aim of modern exposure science is not just to identify exposure for estimating risks, but more importantly, to offer tools that address events in the human body after the toxicant has entered the environment. In the foreseeable future, it is almost impossible to detect the total contaminants in all the exposure pathways. It was shown that heterogeneity in biological measurements is less likely to skew results than heterogeneity in external exposure (Fowler, 2013). Thus, biomonitoring is a sound solution for addressing chemical 'events' and leading to fewer uncertainties. Benefits and limitations of different sample types, including blood, hair, urine or breast milk, have been well summarised (Paustenbach and Galbraith, 2006). Based on the U.S. National Health and Nutrition Examination Survey, a recent study collected information on the presence and concentrations of >400 chemicals in human blood and urine, and then compared their respective hazards by using the biomonitoring equivalents to provide a cross-chemical assessment perspective on the measured levels for approximately 130 analytes (Aylward et al., 2013). Although current developments in analytical approaches and devices have made it possible to detect smaller and smaller concentrations of contaminants, the generalised utility and implications of the biomonitoring information have not as yet been well established.

At least three advantages have been proved when employing biomonitoring to reduce the uncertainties in exposure science. Firstly, in revealing long-term trends for contaminants in the population, in the past decade a bank of studies has provided contaminant concentration trends and indicated the likely environmental implications. For

Table 4
Comparisons of key uncertainty characterisation in current practice and NexGen paradigm.
Crump et al., 2010a; International Programme on Chemical Safety, 2014; Krewski et al., 2014; Lioy, 2010; U.S. EPA, 2014a

	Current practice	NexGen paradigm
Exposure assessment	Using tiered approach (screening level, qualitative, deterministic and advanced statistical analysis) to determine effects of inputs on exposure	Using the mechanistic exposure model, from source-to-dose model system to understand the sources, process from sources to environmental pathways and effects after biological system receives contaminates: largely relies on biomonitoring data and exposomics research
Hazard identification	Expert judgement to weight multiple evidence	*in vitro* testing in human cell lines, QSARs
Dose–response assessment	Determination of PoD: BMD or NOAEL/LOAEL Establishment of Rfd: Uncertainty factor, body scaling, CSAF, PBPK model	Determination of PoD: *in vitro* data or short-term alternative animals Interspecies: integrate data from mammalian species or *in vitro* data for molecular mechanism; comparing molecular epidemiological and clinical studies Intraspecies: system biology and computational toxicology to examine genomic variation, early-life exposures and variability in internal dosimetry Other extrapolation: sensitive assays for environmental dose, route and duration
Risk characterisation	Expert judgement to consider uncertainties in site-specific scenario, toxicity criteria and parameter selection.	Probabilistic approaches characterise overall uncertainties, and determine the significant sources.

Abbreviations: PoD: point of departure; RfD: reference dose; NOAEL: no observed adverse effect level; LOAEL: lowest observed adverse effect level; BMD: benchmark dose; CSAF: chemical-specific adjustment factor; PBPK: Physiologically-based Pharmacokinetic; QSAR: quantitative structure–activity relationship; NexGen: next generation.

example, in Sweden, it has been observed that PFOS and PFHxS levels in composite milk samples were relatively unchanged between 1996 and 2004, with only 20% and 32% coefficients of variation, respectively (Kärrman et al., 2007). However, a remarkable 13-fold increase for PFOS has been observed in females in China (Jin et al., 2007). Secondly, biomonitoring can help develop PBPK models that understand the ADME process. Redding et al. used physiological parameters from a population cohort in Taiwan and reference values given in the literature to estimate partition coefficients based on chemical structure and lipid content in various body tissues (Redding et al., 2008). They also utilised exposure data from Japan to predict the acquired PCB-153 body burdens at an average child-bearing age of 25 years. Good agreement between human biomonitoring measurements and prediction indicated the feasibility of the application of biomonitoring data in HRRA. Thirdly, reconstructing allocation of relevant pathways with advanced statistical techniques may offer solutions for exposure minimisation or even elimination. Long-term external monitoring data are always lacking, and external exposure cannot be determined to completely account for internal exposure (Bernillon and Bois, 2000). Since biomonitoring information is an integrated exposure finding, it offers the opportunity to trace and mimic a realistic exposure scenario (Dong and Hu, 2011; Lyons et al., 2008; Yang et al., 2010). Lyons et al. demonstrated the use of a computational framework that integrates PBPK modelling, BHM and biomonitoring, in order to obtain a population estimate of environmental chloroform source concentrations: the posterior (geometric) standard deviation has fallen by 30% (Lyons et al., 2008). By using Bayesian inference, predicting the biomarker data improved through calibration to chloroform exposure (Yang et al., 2010).

4.2. In vitro data

In the sphere of toxicology, most whole animal data consume resources and furthermore based on high dose the date produces a significantly adverse effect. The toxicity data gap for the large number of chemicals, mixtures and multiple endpoints was recognised by the NAS. In the NRC's version of the conceptual framework for testing toxicity, it has been necessary to shift from testing in animals to *in vitro* assays conducted in human cells to understand the mode of action by which chemicals perturb biological processes (U.S. NRC, 2007). Crump et al. have envisioned the application of *in vitro* data for aiding in PoD determination and extrapolation (Crump et al., 2010a). To infer an *in vivo* human exposure 'standard' using *in vitro* methodologies requires the integration of toxicity pathway modelling and pharmacokinetic modelling (Bhattacharya et al., 2011). Comparing to the BMD or NOAEL/LOAEL approaches to establish the PoD, *in vitro* experiment will be applied to calculate a PoD as a cellular concentration, and the PBPK model can then convert it into a human PoD. For example, Strikwold et al. estimated the PoD as 93.5 mg/kg bw and 26.9 mg/kg bw when using the rat and human cell, respectively, compared to the *in vivo* NOAELs with 333 mg/kg bw for embryo-toxicity of phenol (Strikwold et al., 2013). This study has elucidated the interspecies kinetic differences and contributed to animal testing being reduced, refined and replaced. It has also estimated the ratio between NOAEL and LOAEL and accounted for the uncertainty factor.

Extrapolation across species may be observed at different levels of biological organisation, including cell, organ or individual levels. Based on the similarity of involved pathway, the *in vitro* data appears to be useful for extrapolations (U.S. EPA. 2014a). It has been acknowledged that the toxic effects of TCDD are initiated through the AhR (Fernandez-Salguero et al., 1996; Peterson et al., 1993). Evidence from *in vitro* assays, together with evidence from short-term *in vivo* biomarker responses, and *structure–activity*, shows an order of magnitude or more difference in sensitivity for early-life-stage mortality to TCDD between lake trout and zebrafish, domestic chicken and other avian species (Elonen et al., 1998; Head et al., 2008). Meanwhile, using the human cell lines obtained from normal and sensitivity populations can identify the genomic variability and thus has the potential for interspecies extrapolation. Recent panels including the International HapMap Project and the NCI60 cell panel, have provided a key database of publicly available genotypic, expression, and phenotypic data while allowing researchers to generate their own data related to drug treatment. They can identify genetic variations of interest, and the availability of human lymphoblastoid cell lines from major world populations of European, African, Chinese, and Japanese ancestry has the potential to quantify interindividual and interpopulation differences (Welsh et al., 2009). Compared to animal models, human *in vitro* assays are more reliable since they reflect human physiology and toxicity pathways. The application sensitive assays for environmental dose, route and duration can be used for other types of extrapolations. Case studies include employing an iterative model that optimised the ability of endocrine-related HTS assays to predict components of EDSP T1S and related results at the environmental exposure level (Rotroff et al., 2013).

While a number of advantages can be imagined when employing *in vitro* data, more issues need to be considered. Due to the complexity of the physiological environment, it is widely acknowledged that *in vitro* cell culture systems are poorly representative of human or animal physiology (Lipscomb et al., 2012). For example, when the relationship between cellular response and a downstream apical response poses a challenge, determining PoDs based on *in vitro* data requires more scientific judgement than extrapolating that from *in vivo* data (Crump et al., 2010a). Also, *in vitro* testing will encounter challenges in the form of volatile chemicals, multiple endpoints, metabolism, feedback signalling, and long-term effects.

4.3. Computational toxicology

High content technologies on genome arrays, proteomics, and metabolomics allow us to observe significant changes at many levels, such as mRNAs, proteins and small molecules. When these data streams are classified into pathways and networks, each network may include large numbers of interactions from genes, proteins and metabolites (Fowler, 2013). However, a backlog arising since accumulated data has largely outstripped the interpretation of the data. Facilitating data explanation and understanding these interactions boost the development of computational toxicology. Computational toxicology attempts to synthesise toxicology, statistics, biology and computer modelling have been attempted to predict the toxicity of agents on human health. Compared to traditional testing strategies, it has become possible to identify important biological processes that may be disrupted by chemicals much more efficiently and tracing those biological disruptions to related dose and human exposure to chemicals at many levels.

Computational toxicology began with the development of the QSAR model. The QSAR model is one of the major tools for prioritisation and screening in the context of NexGen HHRA. Traditional QSAR models were established based on chemical descriptors alone, while the predictivity of modern QSAR model was improved by the availability of massive data warehouses (Rusyn and Daston, 2010). Sedykh et al. used published qHTS-response data for 1408 substances in an effort to improve the accuracy of QSAR models, and their results indicated that external classification and coverage of these hybrid models were superior to conventional ones (Sedykh et al., 2011). Zhu et al. have demonstrated the prediction accuracy of the QSAR model did significantly improve (72.7%) when chemical descriptors were augmented by HTS data (Zhu et al., 2008). In current practice, the way a molecule is quantified in terms of descriptors and how the relationship between these chemical descriptors and the toxicological endpoint of interest, constitute the major differences between the various computational methods (Ekins, 2007). OECD has developed three versions of the QSAR toolbox to make this technology readily accessible, transparent, and less demanding in terms of infrastructure costs (Organisation for Economic Co-operation and Development, 2009). Based on integrated information and data from various sources, the current version is able to identify

structural characteristics, potential mechanism of the target chemical, and fill the data gaps by using existing experimental data (Organisation for Economic Co-operation and Development, 2009).

The established database has provided a resource for computational model-building. Warren and Eftim summarised data that is publicly available across a range of biological, chemical and toxicological space (Fowler, 2013). For example, ACToR has accumulated much toxicity data from a variety of peer reviewed studies, and hence for any new chemical, the toxicological information can be sought by its structure or substructure. Rusyn and Daston reviewed the applications of *in vitro* methodologies, toxicogenomics, proteomics, and metabolomics in computational toxicology (Rusyn and Daston, 2010). The prediction of computational toxicology is based on the concept of toxicity pathways. Toxicity pathways offer a causal link between adverse response and action mode. Originating in the era of pathway identification and regulation to pathway quantification, it is crucial to employ computational techniques to underpin the correlations for the biological process among the available data.

In the last decade, numerous studies have established computational models to illustrate cellular, organ or organismal function (Gilchrist et al., 2006; Royle and Kontoravdi, 2013; Steven Wiley et al., 2003). Combining the PBPK model, human biological biomonitoring data and statistical technique is a common method for addressing uncertainties as discussed in the previous section. In the context of NexGen risk assessment, toxicity pathways are required for extrapolations. By studying basic processes in the liver such as antioxidant, bile, energy, and nutrient metabolism to describe liver homeostasis and its perturbed state, Kalyanasundaram has illustrated how to integrate system biology to develop a generic dynamic systems model with *in vitro* measurements, making it possible to predict *in vivo* toxicity (Fowler, 2013). More recently, the AOP concept illustrates a cascade of responses at the macro-molecular, cellular, organ, individual and population levels after the toxicant interacts with a receptor, which provide a framework for uncertainty minimisation (Ankley et al., 2010).

Case studies include how to translate non-mammalian species data and test data to HHRA requirements when the data is insufficient for limited-scope assessments (Fowler, 2013). The prototypes on benzene and other leukemogens, ozone, benzo[a]pyrene/polycyclic aromatic hydrocarbons have also demonstrated that a well described AOP network can help data-limited chemicals based on AOP network similarities for major-scope assessments (Krewski et al., 2014; U.S. EPA, 2014a). Despite advances in computation, challenges still exist regarding how to validate model predictability and various data sources which may lead to risk being misestimated. Additionally, both the standard guidance for computational toxicology and conceptual framework to identify interactions between chemicals have not yet been generalised.

It has long been acknowledged that uncertainty characterisation in HHRA is important. In the past half century, various approaches have been proposed to fill gap between data and regulation requirements. Use of expert judgement, defaults and tools for quantitative uncertainty minimisation attempts to maximise use available resources. The experience of researching TCDD and TCE illustrated how to maximise available information to determine uncertainties, and thereby provide an 'adequate' protection to contaminant exposure. For the NexGen framework, three cornerstones have been summarised (Krewski et al., 2014): 1) toxicity pathway-based approach; 2) a population health approach; and 3) emergence of new risk assessment methodologies. In the context of uncertainty minimisation, the NexGen approach focuses on 'identify most important uncertainty and utilise probabilistic risk assessment to characterise overall uncertainty'. A significant feature of the NexGen's uncertainty reduction is the exploded data across levels from cell to organ, and population, which potentially can enable the mechanism model to meet the 'adequate' safety goal. Advances in molecular biology, omics, computational techniques, and exposure fields are expected to determine, understand and address the chemical behaviour. From another perspective, new challenges may also arise on how to interpret and link the huge data net, such as extrapolation from *in vitro*/*in silico* data to *in vivo* data, and correlations among multiple-level data. In the near future, we can see how both opportunities and challenges from biomonitoring, *in vitro* data and computational toxicology can improve the HHRA and uncertainty minimisation. Although we do not yet comprehend how the exposure science and modern toxicology will develop in the long-term, emerging techniques promise to narrow the gap between available knowledge and regulation requirements.

Acknowledgement

We would like to thank the Cooperative Research Centre for Contamination Assessment and Remediation of the Environment (CRC CARE) for financial support.

References

An, W., Hu, J., Giesy, J.P., Yang, M., 2009. Extinction risk of exploited wild roach (*Rutilus rutilus*) populations due to chemical feminization. Environ Sci Technol 43, 7895–7901.
Andersen, M.E., Krewski, D., 2009. Toxicity testing in the 21st century: bringing the vision to life. Toxicol Sci 107, 324–330.
Ankley, G.T., Bennett, R.S., Erickson, R.J., Hoff, D.J., Hornung, M.W., Johnson, R.D., Mount, D.R., Nichols, J.W., Russom, C.L., Schmieder, P.K., 2010. Adverse outcome pathways: a conceptual framework to support ecotoxicology research and risk assessment. Environ Toxicol Chem 29, 730–741.
Armstrong, B., Hutchinson, E., Unwin, J., Fletcher, T., 2004. Lung cancer risk after exposure to polycyclic aromatic hydrocarbons: a review and meta-analysis. Environ Health Perspect 112, 970–978.
Axelrad, D.A., Bellinger, D.C., Ryan, L.M., Woodruff, T.J., 2007. Dose–response relationship of prenatal mercury exposure and IQ: an integrative analysis of epidemiologic data. Environ Health Perspect 115, 609–615.
Aylward, L.L., Brunet, R.C., Starr, T.B., Carrier, G., Delzell, E., Cheng, H., Beall, C., 2005. Exposure reconstruction for the TCDD-exposed NIOSH cohort using a concentration- and age-dependent model of elimination. Risk Anal 25, 945–956.
Aylward, L.L., Kirman, C.R., Schoeny, R., Portier, C.J., Hays, S.M., 2013. Evaluation of biomonitoring data from the CDC national exposure report in a risk assessment context: perspectives across chemicals. Environ Health Perspect 121, 287–294.
Bernillon, P., Bois, F.Y., 2000. Statistical issues in toxicokinetic modeling: a Bayesian perspective. Environ Health Perspect 108, 883–893.
Bhattacharya, S., Zhang, Q., Carmichael, P.L., Boekelheide, K., Andersen, M.E., 2011. Toxicity testing in the 21st century: defining new risk assessment approaches based on perturbation of intracellular toxicity pathways. PLoS One 6, e20887.
Bokkers, B.G., Slob, W., 2005. A comparison of ratio distributions based on the NOAEL and the benchmark approach for subchronic-to-chronic extrapolation. Toxicol Sci 85, 1033–1040.
Burgess-Herbert, S.L., Euling, S.Y., 2013. Use of comparative genomics approaches to characterize interspecies differences in response to environmental chemicals: challenges, opportunities, and research needs. Toxicol Appl Pharmacol 271, 372–385.
Callahan, M.A., Sexton, K., 2007. If cumulative risk assessment is the answer, what is the question? Environ Health Perspect 115, 799–806.
Cheng, H., Aylward, L., Beall, C., Starr, T.B., Brunet, R.C., Carrier, G., Delzell, E., 2006. TCDD exposure–response analysis and risk assessment. Risk Anal 26, 1059–1071.
Chiu, W.A., Ginsberg, G.L., 2011. Development and evaluation of a harmonized physiologically based pharmacokinetic (PBPK) model for perchloroethylene toxicokinetics in mice, rats, and humans. Toxicol Appl Pharmacol 253, 203–234.
Chiu, W.A., Okino, M.S., Evans, M.V., 2009. Characterizing uncertainty and population variability in the toxicokinetics of trichloroethylene and metabolites in mice, rats, and humans using an updated database, physiologically based pharmacokinetic (PBPK) model, and Bayesian approach. Toxicol Appl Pharmacol 241, 36–60.
Chiu, W.A., Okino, M.S., Lipscomb, J.C., Evans, M.V., 2006. Issues in the pharmacokinetics of trichloroethylene and its metabolites. Environ Health Perspect 114, 1450–1456.
Clark, L.H., Woodrow Setzer, R., Barton, H.A., 2004. Framework for evaluation of physiologically-based pharmacokinetic models for use in safety on risk assessment. Risk Anal 24, 1697–1717.
Cote, I., Anastas, P.T., Birnbaum, L.S., Clark, R.M., Dix, D.J., Edwards, S.W., Preuss, P.W., 2012. Advancing the next generation of health risk assessment. Environ Health Perspect 120, 1499–1502.
Crump, K.S., Canady, R., Kogevinas, M., 2003. Meta-analysis of dioxin cancer dose response for three occupational cohorts. Environ Health Perspect 111, 681.
Crump, K.S., Chen, C., Louis, T.A., 2010a. The future use of in vitro data in risk assessment to set human exposure standards: challenging problems and familiar solutions. Environ Health Perspect 118, 1350–1354.
Crump, K.S., Chiu, W.A., Subramaniam, R.P., 2010b. Issues in using human variability distributions to estimate low-dose risk. Environ Health Perspect 118, 387.
Dong, Z., Hu, J., 2011. Development of lead source-specific exposure standards based on aggregate exposure assessment: Bayesian inversion from biomonitoring information to multipathway exposure. Environ Sci Technol 46, 1144–1152.
Dorne, J., Renwick, A., 2005. The refinement of uncertainty/safety factors in risk assessment by the incorporation of data on toxicokinetic variability in humans. Toxicol Sci 86, 20–26.

Dorne, J., Walton, K., Renwick, A., 2005. Human variability in xenobiotic metabolism and pathway-related uncertainty factors for chemical risk assessment: a review. Food Chem Toxicol 43, 203–216.
Egeghy, P.P., Judson, R., Gangwal, S., Mosher, S., Smith, D., Vail, J., Cohen Hubal, E.A., 2012. The exposure data landscape for manufactured chemicals. Sci Total Environ 414, 159–166.
Egger, M., Smith, G.D., Phillips, A.N., 1997. Meta-analysis: principles and procedures. BMJ 315, 1533–1537.
Ekins, S., 2007. Computational Toxicology: Risk Assessment for Pharmaceutical and Environmental Chemicals. John Wiley & Sons.
Elonen, G.E., Spehar, R.L., Holcombe, G.W., Johnson, R.D., Fernandez, J.D., Erickson, R.J., Tietge, J.E., Cook, P.M., 1998. Comparative toxicity of 2, 3, 7, 8-tetrachlorodibenzo-p-dioxin to seven freshwater fish species during early life-stage development. Environ Toxicol Chem 17, 472–483.
Emond, C., Birnbaum, L.S., DeVito, M.J., 2006. Use of a physiologically based pharmacokinetic model for rats to study the influence of body fat mass and induction of CYP1A2 on the pharmacokinetics of TCDD. Environ Health Perspect 114, 1394–1400.
ENHEALTH, 2012. Environmental Health Risk Assessment — Guidelines for Assessing Human Health Risks From Environmental Hazards. Environmental Health Standing Committee, Canberra, Australia.
European Chemicals Agency. Guidance for human health risk assessment for biocidal active substances and biocidal products. Helsinki, Finland; 2013.
Fernandez-Salguero, P.M., Hilbert, D.M., Rudikoff, S., Ward, J.M., Gonzalez, F.J., 1996. Aryl-hydrocarbon receptor-deficient mice are resistant to 2, 3, 7, 8-tetrachlorodibenzo-p-dioxin-induced toxicity. Toxicol Appl Pharmacol 140, 173–179.
Fowler, B.A., 2013. Computational Toxicology: Methods and Applications for Risk Assessment.
Freireich, E.J., Gehan, E., Rall, D., Schmidt, L., Skipper, H., 1966. Quantitative comparison of toxicity of anticancer agents in mouse, rat, hamster, dog, monkey, and man. Cancer Chemother Rep 50, 219–244.
Gallagher, C.M., Meliker, J.R., 2010. Blood and urine cadmium, blood pressure, and hypertension: a systematic review and meta-analysis. Environ Health Perspect 118, 1676–1684.
Gelman, A., Bois, F., Jiang, J., 1996. Physiological pharmacokinetic analysis using population modeling and informative prior distributions. J Am Stat Assoc 91, 1400–1412.
Gilchrist, M., Thorsson, V., Li, B., Rust, A.G., Korb, M., Kennedy, K., Hai, T., Bolouri, H., Aderem, A., 2006. Systems biology approaches identify ATF3 as a negative regulator of Toll-like receptor 4. Nature 441, 173–178.
Glass, G.V., 1976. Primary, secondary, and meta-analysis of research. Educational Researcher 5, 3–8.
Glorennec, P., Bemrah, N., Tard, A., Robin, A., Le Bot, B., Bard, D., 2007. Probabilistic modeling of young children's overall lead exposure in France: integrated approach for various exposure media. Environ Int 33, 937–945.
Head, J.A., Hahn, M.E., Kennedy, S.W., 2008. Key amino acids in the aryl hydrocarbon receptor predict dioxin sensitivity in avian species. Environ Sci Technol 42, 7535–7541.
Integrated Risk Information System. Methylmercury (MeHg) (CASRN 22967-92-6). in: U.S. Environmental Protection Agency., ed; 2001.
International Programme on Chemical Safety, 1999. Assessing Human Health Risks of Chemicals: Principles for the Assessment of Risk to Human Health from Exposure to Chemicals. World Health Organization, Geneva.
International Programme on Chemical Safety, 2005a. Chemical-specific Adjustment Factors for Interspecies Differences and Human Variability: Guidance Document for Use of Data in Dose/Concentration–Response Asses: IPCS Harmonization Project Document. World Health Organization, Geneva.
International Programme on Chemical Safety. Guidance document on characterizing and communicating uncertainty in exposure assessment. World Health Organization; 2005b.
International Programme on Chemical Safety, 2014. Guidance Document on Evaluating and Expressing Uncertainty in Hazard Characterization. World Health Organization, Geneva.
Jardine, C., Hrudey, S., Shortreed, J., Craig, L., Krewski, D., Furgal, C., McColl, S., 2003. Risk management frameworks for human health and environmental risks. J Toxicol Environ Health B Crit Rev 6, 569–718.
Jin, Y., Saito, N., Harada, K.H., Inoue, K., Koizumi, A., 2007. Historical trends in human serum levels of perfluorooctanoate and perfluorooctane sulfonate in Shenyang. China The Tohoku Journal of Experimental Medicine 212, 63–70.
Johnson, J.B., Omland, K.S., 2004. Model selection in ecology and evolution. Trends Ecol Evol 19, 101–108.
Junko, N., Isamu, O., 2008. Specific Health Risk Assessment for PCB (in Japanese). Maruzen Company Limited, Tokyo.
Kärrman, A., Ericson, I., van Bavel, B., Darnerud, P.O., Aune, M., Glynn, A., Lignell, S., Lindström, G., 2007. Exposure of perfluorinated chemicals through lactation: levels of matched human milk and serum and a temporal trend, 1996–2004, in Sweden. Environ Health Perspect 115, 226–230.
Keil, D.E., Peden-Adams, M.M., Wallace, S., Ruiz, P., Gilkeson, G.S., 2009. Assessment of trichloroethylene (TCE) exposure in murine strains genetically-prone and non-prone to develop autoimmune disease. Journal of Environmental Science and Health Part A 44, 443–453.
Krewski, D., Westphal, M., Andersen, M.E., Paoli, G.M., Chiu, W.A., Al-Zoughool, M., Croteau, M.C., Burgoon, L.D., Cote, I., 2014. A framework for the next generation of risk science. Environ Health Perspect 122, 796–805.
Lanphear, B.P., Hornung, R., Khoury, J., Yolton, K., Baghurst, P., Bellinger, D.C., Canfield, R.L., Dietrich, K.N., Bornschein, R., Greene, T., 2005. Low-level environmental lead exposure and children's intellectual function: an international pooled analysis. Environ Health Perspect 113, 894–899.
Lehman AJ, F.O. 100-fold margin of safety. Associate Food Drug Office US Quarterly Bull. 18:33–35; 1954
Lioy, P.J., 2008. Time for a change: from exposure assessment to exposure science. Environ Health Perspect 116, A282.
Lioy, P.J., 2010. Exposure science: a view of the past and milestones for the future. Environ Health Perspect 118, 1081–1090.
Lioy, P.J., Smith, K.R., 2013. A discussion of exposure science in the 21st century: a vision and a strategy. Environ Health Perspect 121, 405.
Lipscomb, J.C., Haddad, S., Poet, T., Krishnan, K., 2012. Physiologically-based pharmacokinetic (PBPK) models in toxicity testing and risk assessment. New Technologies for Toxicity Testing. Springer US, pp. 76–95.
Lyons, M.A., Yang, R.S.H., Mayeno, A.N., Reisfeld, B., 2008. Computational toxicology of chloroform: Reverse dosimetry using Bayesian inference, Markov chain Monte Carlo simulation, and human biomonitoring data. Environ Health Perspect 116, 1040–1046.
National Toxicology Program. NTP Technical Report on the carcinogenesis bioassay of bisphenol A (CAS No. 80-05-7) in F344 rats and B6C3F1 mice (feed study). NTP-80-35. NIH Publ. No. 82–177. National Toxicology Program; 1982
Organisation for Economic Co-operation and Development. Guidance document for using the OECD(Q) SAR application toolbox to develop chemical categories according to the OECD guidance on grouping of chemicals. ENV/JM/MONO(2009)5 2009
Organisation for Economic Co-operation and Development. Statistical and dose response analysis, including benchmark dose and linear extrapolation, NOAELS and NOAEL, LOAELS and LOELS. 2010.
Paustenbach, D., Galbraith, D., 2006. Biomonitoring and biomarkers: exposure assessment will never be the same. Environ Health Perspect 114, 1143.
Peterson, R.E., Theobald, H.M., Kimmel, G.L., 1993. Developmental and reproductive toxicity of dioxins and related compounds: cross-species comparisons. CRC Crit Rev Toxicol 23, 283–335.
Poulter, S.R. Monte Carlo simulation in environmental risk assessment—science, policy and legal issues. A Risk Assessment and Policy Association; 1998
Redding, L.E., Sohn, M.D., McKone, T.E., Chen, J.-W., Wang, S.-L., Hsieh, D.P., Yang, R.S., 2008. Population physiologically based pharmacokinetic modeling for the human lactational transfer of PCB-153 with consideration of worldwide human biomonitoring results. Environ Health Perspect 116, 1629–1635.
Renwick, A., 1993. Data derived safety factors for the evaluation of food additives and environmental contaminants. Food Addit Contam 10, 275–305.
Renwick, A., Lazarus, N., 1998. Human variability and noncancer risk assessment—an analysis of the default uncertainty factor. Regul Toxicol Pharmacol 27, 3–20.
Rhomberg, L.R., Goodman, J.E., Haber, L.T., Dourson, M., Andersen, M.E., Klaunig, J.E., Meek, B., Price, P.S., McClellan, R.O., Cohen, S.M., 2011. Linear low-dose extrapolation for noncancer health effects is the exception, not the rule. Crit Rev Toxicol 41, 1–19.
Rotroff, D.M., Dix, D.J., Houck, K.A., Knudsen, T.B., Martin, M.T., McLaurin, K.W., Reif, D.M., Crofton, K.M., Singh, A.V., Xia, M., 2013. Using in vitro high throughput screening assays to identify potential endocrine-disrupting chemicals. Environ Health Perspect 121, 7–14.
Royle, K., Kontoravdi, C., 2013. A systems biology approach to optimising hosts for industrial protein production. Biotechnol Lett 35, 1961–1969.
Rusyn, I., Daston, G.P., 2010. Computational toxicology: realizing the promise of the toxicity testing in the 21st century. Environ Health Perspect 118, 1047–1050.
Schultz, T.W., Cronin, M.T., Walker, J.D., Aptula, A.O., 2003. Quantitative structure–activity relationships (QSARs) in toxicology: a historical perspective. J Mol Struct (THEOCHEM) 622, 1–22.
Sedykh, A., Zhu, H., Tang, H., Zhang, L.Y., Richard, A., Rusyn, I., Tropsha, A., 2011. Use of in vitro HTS-derived concentration–response data as biological descriptors improves the accuracy of QSAR models of in vivo toxicity. Environ Health Perspect 119, 364–370.
Starr, T.B., 2001. Significant shortcomings of the US Environmental Protection Agency's latest draft risk characterization for dioxin-like compounds. Toxicol Sci 64, 7–13.
Steenland, K., Deddens, J., Piacitelli, L., 2001. Risk assessment for 2, 3, 7, 8-tetrachlorodibenzo-p-dioxin (TCDD) based on an epidemiologic study. Am J Epidemiol 154, 451–458.
Steven Wiley, H., Shvartsman, S.Y., Lauffenburger, D.A., 2003. Computational modeling of the EGF-receptor system: a paradigm for systems biology. Trends Cell Biol 13, 43–50.
Strikwold, M., Spenkelink, B., Woutersen, R.A., Rietjens, I.M., Punt, A., 2013. Combining in vitro embryotoxicity data with physiologically based kinetic (PBK) modelling to define in vivo dose–response curves for developmental toxicity of phenol in rat and human. Arch Toxicol 87, 1709–1723.
Thornton, A., Lee, P., 2000. Publication bias in meta-analysis: its causes and consequences. J Clin Epidemiol 53, 207–216.
Tice, R.R., Austin, C.P., Kavlock, R.J., Bucher, J.R., 2013. Improving the human hazard characterization of chemicals: a Tox21 update. Environ Health Perspect 121, 756–765.
U.K. IGHRC, 2003. Uncertainty Factors: Their Use in Human Health Risk Assessment by UK Government. Interdepartmental Group on Health Risks from Chemicals, Leicester, U.K.
U.S. EPA, 1986. Guideline for Carcinogen Risk Assessment. 630/R-00/004. U.S. Environmental Protection Agency, Washington, DC.
U.S. EPA. Guidelines for exposure assessment; notice. Part VI. Federal Register 29 May 1992 57(104): 22888–22938. Washington, DC: U.S. Environmental Protection Agency; 1992
U.S. EPA, 1995. The Use of the Benchmark Dose Approach in Health Risk Assessment. U.S. Environmental Protection Agency, Washington, DC.
U.S. EPA, 1996. Guidelines for Carcinogen Risk Assessment. U.S. Environmental Protection Agency, Washington, DC.
U.S. EPA, 1997. Guiding Principles for Monte Carlo Analysis. U.S. Environmental Protection Agency, Washington, DC.
U.S. EPA, 2000. Risk Characterization Handbook. U.S. Environmental Protection Agency, Washington, DC.

U.S. EPA, 2001a. Risk Assessment Guidance for Superfund (RAGS) Volume iii — Part A: Process for Conducting Probabilistic Risk Assessment. U.S. Environmental Protection Agency, Washington DC.
U.S. EPA, 2001b. Value-of-information Approach to Motivate Uncertainty Factors With Mechanistic Data: Chlorine Human Health Inhalation Risk Case Study. U.S. Environmental Protection Agency, Washington, DC.
U.S. EPA, 2002. Guidance on Cumulative Risk Assessment of Pesticide Chemicals That Have a Common Mechanism of Toxicity. U.S. Environmental Protection Agency, Washington, DC.
U.S. EPA, 2003. Exposure and Human Health Reassessment of 2,3,7,8-Tetrachlorodibenzo-p-dioxin (TCDD) and Related Compounds National Academy Sciences (NAS) Review Draft. U.S. Environmental Protection Agency, Washington, DC.
U.S. EPA, 2005. Guidelines for Carcinogen Risk Assessment. U.S. Environmental Protection Agency, Washington, DC.
U.S. EPA, 2006. Approaches for the Application of Physiologically Based Pharmacokinetic (PBPK) Models and Supporting Data in Risk Assessment. U.S. Environmental Protection Agency, Washington, DC.
U.S. EPA, 2009. U.S. EPA Office of Research and Development Computational Toxicology Research Program Implementation Plan for Fiscal Years 2009 to 2012. U.S. Environmental Protection Agency, Washington, DC.
U.S. EPA, 2010. EPA's Reanalysis of Key Issues Related to Dioxin Toxicity and Response to NAS Comments. U.S. Environmental Protection Agency, Washington, DC.
U.S. EPA, 2012a. EPA's Reanalysis of Key Issues Related to Dioxin Toxicity and Response to NAS Comments Volume 1. U.S. Environmental Protection Agency, Washington, DC.
U.S. EPA, 2012b. Fact Sheet: Chemicals Snapshot. U.S. Environmental Protection Agency, Washington, DC.
U.S. EPA, 1985. Health Effects Assessment Document for Polychlorinated Dibenzo-p-dioxins. Environmental Protection Agency, Washington, DC.
U.S. EPA, 1988. Special Report on Ingested Inorganic Arsenic Skin Cancer; Nutritional Essentiality. U.S. Environmental Protection Agency, Washington, DC.
U.S. EPA. Recommended use of body weight 3/4 as the default method in derivation of the oral reference dose. EPA/100/R11/0001. Washington, D.C. Environmental Protection Agency; 2011.
U.S. EPA, 2014a. Next Generation Risk Assessment: Incorporation of Recent Advances in Molecular, Computational, and Systems Biology. Washington, DC.
U.S. EPA, 2014b. Review of EPA's Integrated Risk Information System (IRIS) Process: The National Academies Press.
U.S. NRC, 1983. Risk Assessment in the Federal Government: Managing the Process. U.S. National Research Council, Washington, DC.
U.S. NRC, 1994. Science and Judgment in Risk Assessment. U.S. National Research Council, Washington, DC.
U.S. NRC, 1996. Understanding Risk: Informing Decisions in a Democratic Society. U.S. National Research Council, Washington, DC.
U.S. NRC, 2007. Toxicity Testing in the 21st Century: A Vision and Strategy. U.S. National Research Council, Washington, DC.
U.S. NRC, 2009. Science and Decisions: Advancing Risk Assessment. The National Academies Press, Washington, D.C.
U.S. NRC, 2013. Environmental Decisions in the Face of Uncertainty. The National Academies Press.
Van der Molen, G.W., Kooijman, B., Wittsiepe, J., Schrey, P., Flesch-Janys, D., Slob, W., 1999. Estimation of dioxin and furan elimination rates with a pharmacokinetic model. J Expo Anal Environ Epidemiol 10, 579–585.
Viberg, H., Fredriksson, A., Eriksson, P., 2004. Investigations of strain and/or gender differences in developmental neurotoxic effects of polybrominated diphenyl ethers in mice. Toxicol Sci 81, 344–353.
Wambaugh, J.F., Setzer, R.W., Reif, D.M., Gangwal, S., Mitchell-Blackwood, J., Arnot, J.A., Joliet, O., Frame, A., Rabinowitz, J., Knudsen, T.B., Judson, R.S., Egeghy, P., Vallero, D., Hubal, E.A.C., 2013. High-throughput models for exposure-based chemical prioritization in the expo cast project. Environ Sci Technol 47, 8479–8488.
Wan, Y., Zhang, K., Dong, Z., Hu, J., 2013. Distribution is a major factor affecting bioaccumulation of decabrominated diphenyl ether: Chinese sturgeon (*Acipenser sinensis*) as an example. Environ Sci Technol 47, 2279–2286.
Welsh, M., Mangravite, L., Medina, M.W., Tantisira, K., Zhang, W., Huang, R.S., McLeod, H., Dolan, M.E., 2009. Pharmacogenomic discovery using cell-based models. Pharmacol Rev 61, 413–429.
Whitehead, A., 2003. Meta-analysis of Controlled Clinical Trials. Wiley.
Williams, P., B. Hubbell, E. J. Weber, C. Fehrenbacher, D. Hrdy, V. G. Zartarian. An overview of exposure assessment models used by the U.S. Environmental Protection Agency. Hertfordshire, UK; 2010.
Williams, P.R., Paustenbach, D.J., 2002. Risk characterization: principles and practice. J Toxicol Environ Health B Crit Rev 5, 337–406.
Xu, T.; White, L.; Hui, D.; Luo, Y. Probabilistic inversion of a terrestrial ecosystem model: analysis of uncertainty in parameter estimation and model prediction. Global Biogeochemical Cycles. 20; 2006
Yang, Y., Xu, X., Georgopoulos, P.G., 2010. A Bayesian population PBPK model for multiroute chloroform exposure. J Expo Sci Environ Epidemiol 20, 326–341.
Zeise, L., Bois, F.Y., Chiu, W.A., Hattis, D., Rusyn, I., Guyton, K.Z., 2013. Addressing human variability in next-generation human health risk assessments of environmental chemicals. Environ Health Perspect 121, 23–31.
Zhu, H., Rusyn, I., Richard, A., Tropsha, A., 2008. Use of cell viability assay data improves the prediction accuracy of conventional quantitative structure–activity relationship models of animal carcinogenicity. Environ Health Perspect 116, 506–513.