Regulatory Toxicology and Pharmacology 146 (2024) 105525

Contents lists available at ScienceDirect

Regulatory Toxicology and Pharmacology

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

The 2022 world health organization reevaluation of human and mammalian toxic equivalency factors for polychlorinated dioxins, dibenzofurans and biphenyls

Michael DeVito [a,1], Bas Bokkers [b], Majorie B.M. van Duursen [c], Karin van Ede [d], Mark Feeley [e], Elsa Antunes Fernandes Gáspár [d], Laurie Haws [f], Sean Kennedy [g], Richard E. Peterson [h], Ron Hoogenboom [i], Keiko Nohara [j], Kim Petersen [k,*], Cynthia Rider [l], Martin Rose [m,n], Stephen Safe [o], Dieter Schrenk [p], Matthew W. Wheeler [q], Daniele S. Wikoff [r], Bin Zhao [s,t], Martin van den Berg [u]

[a] *Center for Computational Toxicology and Exposure, United States Environmental Protection Agency, Research Triangle Park, NC, USA*
[b] *Centre for Safety of Substances and Products, National Institute for Public Health, And the Environment (RIVM), Bilthoven, the Netherlands*
[c] *Amsterdam Institute for Life and Environment, Environmental Health & Toxicology, Vrije Universiteit, Amsterdam, the Netherlands*
[d] *KeyToxicology, Arnhem, the Netherlands*
[e] *Mark Feeley, Ottawa, Canada*
[f] *ToxStrategies, Inc., Austin, TX, USA*
[g] *Department of Biology, University of Ottawa, Canada*
[h] *School of Pharmacy/University of Wisconsin, Madison, WI, USA*
[i] *Wageningen Food Safety Research (WFSR), Wageningen, the Netherlands*
[j] *Health and Environmental Risk Division, National Institute for Environmental Studies, Tsukuba, 305-8506, Japan*
[k] *Department of Nutrition and Food Safety, Standards and Scientific Advice on Food and Nutrition, World Health Organization, Geneva Switzerland*
[l] *National Institute of Environmental Health Science, Division of the Translational Toxicology, Durham, USA*
[m] *FERA Science Ltd, Sand Hutton, York, YO41 1LZ, UK*
[n] *Manchester Institute of Biotechnology, University of Manchester, Manchester, UK*
[o] *Department of Veterinary Physiology and Pharmacology, Texas A&M University, College Station, TX, 77843, USA*
[p] *Food Chemistry and Toxicology Department, University of Kaiserslautern, D-67663, Kaiserslautern, Germany*
[q] *Biostatistics and Computational Biology Branch, National Institute of Environmental Health Sciences, RTP, NC, USA*
[r] *ToxStrategies, Inc., Asheville, NC, USA*
[s] *State Key Laboratory of Environmental Chemistry and Ecotoxicology, Research Center for Eco-Environmental Sciences, Chinese Academy of Sciences, Beijing, 100085, China*
[t] *University of Chinese Academy of Sciences, Beijing, 100049, China*
[u] *Institute for Risk Assessment Sciences, Utrecht University, Yalelaan 104, 3584 CM, Utrecht, the Netherlands*

ARTICLE INFO

Handling Editor: Lesa Aylward

ABSTRACT

In October 2022, the World Health Organization (WHO) convened an expert panel in Lisbon, Portugal in which the 2005 WHO TEFs for chlorinated dioxin-like compounds were reevaluated. In contrast to earlier panels that employed expert judgement and consensus-based assignment of TEF values, the present effort employed an update to the 2006 REP database, a consensus-based weighting scheme, a Bayesian dose response modeling and meta-analysis to derive "Best-Estimate" TEFs. The updated database contains almost double the number of datasets from the earlier version and includes metadata that informs the weighting scheme. The Bayesian analysis of this dataset results in an unbiased quantitative assessment of the congener-specific potencies with uncertainty estimates. The "Best-Estimate" TEF derived from the model was used to assign 2022 WHO-TEFs for almost all congeners and these values were not rounded to half-logs as was done previously. The exception was for the mono-ortho PCBs, for which the panel agreed to retain their 2005 WHO-TEFs due to limited and

* Corresponding author.
E-mail address: kpetersen@who.int (K. Petersen).

[1] This work is not a product of the United States Government or the U.S. Environmental Protection Agency. MdV is not doing this work in any governmental capacity. The views expressed are his own and do not necessarily represent those of the United States or U.S. EPA.

https://doi.org/10.1016/j.yrtph.2023.105525
Received 26 July 2023; Received in revised form 21 October 2023; Accepted 1 November 2023
Available online 14 November 2023

heterogenous data available for these compounds. Applying these new TEFs to a limited set of dioxin-like chemical concentrations measured in human milk and seafood indicates that the total toxic equivalents will tend to be lower than when using the 2005 TEFs.

1. Introduction

Chlorinated dioxins and dioxin-like compounds (DLC's) refer to the family of structurally and toxicologically related polychlorinated dibenzo-p-dioxins (PCDDs), polychlorinated dibenzofurans (PCDFs) and dioxin-like polychlorinated biphenyls (PCBs). Human background exposure to these DLC's is primarily through the diet, with food of animal origin being the most important source (USEPA 2012, Knutsen, Alexander et al., 2018). Strict regulatory controls on major industrial sources and regulatory national monitoring programs that screen and quantify the presence of PCDDs, PCDFs and PCBs in feed and food have contributed to a significant reduction in human exposure by approximately 90% since the 1960s (Knutsen, Alexander et al., 2018; Hays and Aylward 2003). As a result, a significant global decline in plasma and human milk levels has been observed in the general population during the last decades (Hays and Aylward 2003, Muzembo et al., 2019). Human exposure and systemic levels are considerably influenced by external factors such as low or high consumption of animal products, living in an industrialized area or not and age (Hays and Aylward 2003; USEPA 2012, Knutsen, Alexander et al., 2018). Despite substantial reductions in human exposure and associated body levels, there is still concern from a toxicological point of view that present exposures are still above those considered safe for human health (Knutsen, Alexander et al., 2018).

Decades of toxicological and mechanistic research demonstrate that most, if not all, adverse health effects associated with exposure to DLC's are mediated through the aryl hydrocarbon receptor (AhR). This AhR is a ligand-activated nuclear transcription factor that is present in many cells of most modern bilaterian animal species (Hahn et al., 2017). In absence of a ligand, the AhR is present in the cytosol as a multiprotein complex containing a heat shock protein 90 (hsp 90), the HBV X-associated protein (XAP2), and the co-chaperone protein p23 (Hankinson 1995; Beischlag et al., 2008). AhR translocation into the nucleus and the subsequent interaction with DNA response elements can be initiated by binding of both exogenous and endogenous chemicals to the AhR and altering the expression of various genes, which results in a diverse spectrum of biological and toxicological effects (White and Birnbaum, 2009; Denison et al., 2011). The functional characteristics of the AhR are broadly conserved among vertebrate species and upon AhR activation by DLC's a wide variety of species-specific toxic and biological effects has been reported (Denison et al., 2011; Denison and Faber 2017). These effects are especially observed for those DLC's that have a halogenated substitution pattern on the four lateral positions, which are numbered 2, 3,7,8 in dibenzo-*p*-dioxins or dibenzofurans and 3,4 in biphenyls. As a result, those "planar" DLC's are considered the most toxicologically relevant for human risk assessment (Safe 1990).

The biochemical and toxic responses have been studied thoroughly in experimental animals for many decades and are characterized by enzyme induction, retinoid changes, severe weight loss, thymic atrophy, hepatotoxicity, immunotoxicity, endocrine disruption, reproductive and developmental effects, and tumorigenesis (Safe 1990; Birnbaum 1994; Birnbaum and Tuomisto 2000). In humans, long term exposure to 2,3,7,8-tetrachlorodibenzo-p-dioxin (2,3,7,8-TCDD), the most potent and best studied DLC, is linked to impairment of the immune system, developing nervous system, the endocrine system, reproductive functions, and carcinogenic responses (JEFCA, 2002; Loomis et al., 2018; USEPA 2012, Knutsen, Alexander et al., 2018).

Assessing the potential health risks associated with exposure to DLC's remains challenging, as humans and wildlife are exposed to complex mixtures (Safe 1994) from which the 2,3,7,8 substituted PCDDs and PCDFs have the highest tissue retention (van den Berg et al., 1989). For PCBs the tissue retention in humans is more complicated, as both planar (3,4 and 3′,4′-substituted) as well as non-planar (2 substituted) congeners can be found in human samples (e.g., blood, milk, and adipose tissue). Based on the observation that DLC's share the same mechanism of action, supported by a broad range of in vitro and in vivo experiments, it is generally accepted that congener specific toxicities are additive, but their relative potencies differ (Safe 1990; Safe 1994; USEPA 2010, Chain, Knutsen et al., 2018; Van den Berg, Birnbaum et al. 1998; Van den Berg, Birnbaum et al. 2006). This led to the development of the toxic equivalency concept in which each congener has a specific toxic equivalency factor (TEF) calculated by comparing the potency of the congener to that of the reference compound, 2,3,7,8-TCDD (Safe 1990; Safe 1994, Van den Berg, Birnbaum et al. 1998; Van den Berg, Birnbaum et al. 2006). These congener specific TEF values are multiplied by their respective congener concentrations and the resultant products are summed to derive a 2,3,7,8-TCDD toxic equivalence (TEQ) of the mixture in a sample, e.g., a food product or human milk. This TEQ value of the mixture is assumed to behave as an equivalent concentration of 2,3,7,8-TCDD and used to estimate potential health risks compared to regulatory toxicity values (Reference Dose or Tolerable Weekly Intake) for 2,3,7,8-TCDD. The TEF methodology is also used in epidemiological studies to relate dioxin exposure to adverse effects such as semen quality (Knutsen et al., 2018) in the setting of regulatory toxicity values for 2,3,7,8-TCDD. This TEQ concept is used to support world-wide risk characterization of DLC's in food and feed and has also been used to assess exposure to DLC's in human tissues (Van den Berg, Birnbaum et al. 2006; USEPA 2010; USEPA 2012, Knutsen, Alexander et al., 2018).

The WHO has held several expert meetings evaluating the TEF methodology and assigning TEF values to DLC's beginning in the mid 1990's (Ahlborg and Hanberg 1994; Birnbaum 1994, Van den Berg, Birnbaum et al. 1998, Van den Berg, Birnbaum et al. 2006). The recommended WHO-TEF values coming out of these expert meetings have been adopted by regulatory agencies world-wide. Since the last WHO expert meeting in 2005, a large amount of new data has been published that could provide more accurate estimates of congener specific relative effect potencies (REPs) and derivation of TEFs. The increasing wealth of experimental data now also allows a more detailed determination of the uncertainty surrounding these REPs. For the 2005 WHO expert meeting an extensive REP database was developed by Haws et al., (2006); Haws et al., 2006) that supported the derivation of the 2005 WHO-TEF values. For the current review, an updated database (Fitch et al., submitted) was generated which incorporates many new studies with DLC's comprising over 700 additional REP data sets, which were not available during the 1998 and 2005 WHO expert meetings (Van den Berg, Birnbaum et al. 1998, Haws et al., 2006, Van den Berg, Birnbaum et al. 2006.

Previous WHO expert meetings relied heavily on expert judgment when deriving TEF values from an underlying database of REPs. However, during the last decade significant advances have been made in quantitative approaches that better allow for use of all available data. Such computational and statistical approaches allow for integration of reliability concepts and allows for the analysis of dose-response relationships using machine-learning and Bayesian meta-regression methodology, respectively. These types of methods are becoming accepted practices when analyzing toxicological data. The availability of an expanded REP database, a consensus-based weighting framework, and recent acceptance of Bayesian methodology led the WHO to convene an expert meeting in Lisbon (Portugal) on October 17 to 21, 2022. During this meeting new data on the relative potency of DLC's

compared to TCDD were reviewed, and the Bayesian method was used to derive updated TEFs and surrounding uncertainties, for comparison to the 2005 WHO TEF values. During this meeting several key issues on the general concepts related to TEF methodology (e.g., intake vs. systemic REPs, reliability of individual study types, commonality of biological pathways) were also discussed. Assessing, the impact of the 2022 WHO TEFs on future risk assessments and evaluation of brominated analogs were not within the scope of this meeting. Of note is that the toxicity of TCDD and DLC's in humans was not a subject of this meeting. This article presents the consensus findings of this meeting and recommendations of the WHO to update the 2005 WHO-TEFs (Van den Berg, Birnbaum et al. 2006).

2. Approach and process of the expert meeting

The WHO announced a call for experts on May 6, 2022, with the criteria and process for selection of experts accompanying the call for the "WHO initiative to update the 2005 WHO-TEF for dioxin and dioxin-like compounds".[2] The objective of the call was to identify qualified independent experts willing to serve as a member of the expert panel. This panel was charged to review new relative potency data for DLC's and consider the possible need to update the 2005 WHO-TEFs. In addition, this panel was asked to advise WHO on the Bayesian methodology and, if needed, to update the 2005 WHO-TEF values for DLC's.

Preceding this 2022 expert meeting, the WHO has worked for the last two years with a small task force of internationally recognized independent experts to lay the groundwork for this meeting. Based on the recommendations of this task force, the European Food Safety Authority (EFSA) engaged with two consulting firms, ToxStrategies (US-based) and KeyToxicology (NL-based). Each consultancy firm had a separate scope of work to be completed for EFSA that was related to preparing materials for the 2022 WHO meeting. These efforts focused on updating, refining, and reviewing the TEF database as established earlier (Haws et al., 2006), developing and reviewing statistical approaches with special focus on the Bayesian methodology, quantitative analysis of the database, and jointly presented these findings at the WHO expert meeting. Following review and comments from the workgroup, these presentations have subsequently been published (Ring et al., 2023; Wikoff et al., 2023; Fitch et al., 2023).

3. Development of the REP database

Over several decades, the REP database has significantly evolved with respect to content, volume, accuracy, level of QC review, and sophistication. The original database, prepared at the Karolinska Institute, was an Excel-based collection of individual REP values taken directly from the literature and used for the 1994 (Ahlborg and Hanberg 1994) and 1998 (Van den Berg, Birnbaum et al. 1998) WHO TEF evaluations. In preparation for the 2005 WHO expert meeting ToxStrategies collaborated with several external partners to develop a refined database of relative potency estimates (Haws et al., 2006). Besides being a significant extension of the Karolinska database, with new studies, the refined database also incorporated study inclusion and exclusion criteria and established a transparent record concerning database curation (Haws et al., 2006).

Following the 2005 WHO TEF meeting, ToxStrategies continued working with external collaborators on updating the REP database through 2021 and developing a consensus-based REP weighting methodology. To accommodate the key scientific issues concerning quality and REP derivation approaches, the database update was done using formal systematic review methods. This included development of a protocol a priori, which was provided to EFSA for review and approval (Fitch et al., 2023).

To be included in the database, studies needed to meet the following criteria (for full inclusion/exclusion, see Fitch et al., 2023).

- *Dose response data for a reference compound (i.e., TCDD or PCB126) and at least one DLC. Studies examining mixtures of DLC's were excluded as these type studies cannot be used to determine congener specific differences in biochemical or toxic effects.*
- *Statistically significant responses compared to an untreated or vehicle control were observed for both a reference compound and at least one DLC.*
- *Experimental systems included non-marine mammals, a mammalian cell line, or cells transfected with a relevant sequence (e.g., DR CALUX).*
- *Only peer-reviewed publications reporting original study data were included. To avoid duplication of data, conference abstracts, and reviews were excluded.*

The development of the latest version of the database utilized DistillerSR (Ottawa, CN) as the platform for screening and extracting data; a relational database was necessary (vs., an Excel spreadsheet) because of the increased amount of information collected for each study. Of note, generation of this database included both the addition of information from new studies, as well as collection of dose-response information from studies in the previous database. The development of the database, including the workflow and study evaluation criteria, are more fully described in (Fitch et al., 2023; Ring et al., 2023; Wikoff, 2023). As a result, the present database almost doubled in size since 2005 and now includes more than 700 additional congener specific dose response REP datasets. The database underwent 100% internal quality control review of all extracted data. Subsequently, ToxStrategies was contracted by the EFSA (EFSA Contract - NP/EFSA/SCER/2021/01) to provide this updated database to the WHO and publish this revised database in the peer-reviewed literature (Fitch et al., 2023). EFSA also awarded a contract to KeyToxicology to conduct a peer review of this ToxStrategies database. KeyToxicology focused on a selected subset of congeners (1,2,3,4,6,7,8-HpCDD, 2,3,4,7,8-PeCDF, 1,2,3,7,8-HxCDF, 1,2,3,6,7,8 HxCDF, PCB-126, and PCB -169), covering approximately 50% of the newly added datasets. These congeners are responsible for approximately 80–90% of the total TEQ (calculated using the 2005 WHO TEFs) in feed, food, and human tissue. While not all extracted datasets were reviewed by KeyToxicology, the review covered 338 newly added REP datasets which were considered representative of the newly extracted data.

The KeyToxicology review of the database found that newly added studies for the selected congeners followed the 2005 WHO-TEF criteria (Van den Berg, Birnbaum et al. 2006). Further, KeyToxicology observed only very minor discrepancies in the database when compared with the original studies. KeyToxicology concluded that there were no significant and/or systematic problems within the ToxStrategies database. Moreover, all minor changes and clarifications provided by KeyToxicology were accepted by ToxStrategies and included into the database. It should be noted that the previous REP databases used to derive WHO TEFs did not have this level of rigorous peer-review. As a result, this effort and review process has largely increased the confidence in the REP database that the 2022 WHO expert panel used in making decisions regarding changes to any of the 2005 WHO-TEFs.

4. Weighting scheme

In previous WHO TEF evaluations, a qualitative weighting scheme was based on expert judgment that prioritized data with the following schema: chronic toxicity data > subchronic toxicity data > acute toxicity > biochemical responses>in vitro > QSAR results (Van den Berg, Birnbaum et al. 1998; Van den Berg, Birnbaum et al. 2006). Following the 2005 WHO TEF review, a multiyear effort was undertaken to develop and implement a quantitative weighting framework (Wikoff

[2] https://www.who.int/news-room/articles-detail/call-for-experts-who-initiative-to-update-the-2005-who-tef-for-dioxin-and-dioxin-like-compounds.

et al., submitted). REP estimates were weighted in a manner consistent with the informal weighting procedure used by previous WHO expert panels.[3] The panel, using a consensus-based approach, identified six main-study characteristics most impactful in weighting an individual REP for human health risk assessment.

1. Study type.
2. Study model.
3. Pharmacokinetics.
4. REP derivation quality.
5. Endpoint
6. REP derivation method.

Study type was described as either in vivo or in vitro, with in vitro further described as human primary, human immortalized, non-human mammalian cell lines (Wikoff, 2023). For the in vivo studies, different species, strains, and gender were given equal weight. There were no human in vivo relative potency data included in the database. Study model describes the level of biological complexity from organismal to unicellular. For pharmacokinetics, if the test congener had similar pharmacokinetic properties to TCDD or PCB126, then exposure paradigms, in vivo or in vitro, were not a critical factor. However, if the half-life or bioavailability were significantly different than TCDD or PCB126, then study duration was important to consider in study quality. REP derivation quality focused on the number of dose levels, sample size per dose group and if the test congener and reference congener attained a maximum response. The endpoint category was consistent with previous TEF panels in that toxic responses were weighted more than biochemical which were weighted more than QSAR. (Wikoff et al., 2023). REP Derivation Method divided studies into high, medium, low and QSAR quality categories. The high category statistically modeled the REP and accounted for parallelism. The medium category statistically modeled the data but did not account for parallelism. The low category employed NOEL/LOEL ratios or response ratios. The endpoint category was consistent with previous TEF panels in that toxic responses were weighted more than biochemical which were weighted more than QSAR (Wikoff et al., submitted).

5. Derivation of bayesian TEF estimates from weighted evaluations of REP dose-response data

The 2005 WHO TEF panel identified two key challenges in using underlying data to directly assign.

- *a large range in study quality.*
- *the number of different methods used to derive REPs.*

In response to these challenges, over the course of several years, ToxStrategies and their collaborators developed both a weighting framework and a workflow to derive a method to determine the Bayesian estimate of a TEF. As part of this effort, this group of scientists introduced the term "Best-Estimate TEF" or BE-TEF. This is described in detail by Ring et al., (Ring et al., 2023). After the appropriate experimental data were collected, the developed workflow comprised of four distinct sections.

- *Machine-learning based quality weighting of REPs.*
- *Bayesian dose-response modeling.*
- *Bayesian meta-analysis.*
- *Derivation of the congener-specific BE-TEFs.*

In view of the methodology underlying the Bayesian methodology, the pre-meeting WHO task force decided to apply an independent peer review by Dr. Matthew Wheeler to this method and applied the workflow by Dr. Wheeler prior to the October 2022 expert meeting. All recommendations coming from this peer review regarding the analyses and models employed were incorporated by ToxStrategies prior to the WHO expert meeting and are further described in Ring et al., submitted (Ring et al., 2023). The methodology presented in short below has been described in detail by (Fitch et al., 2023; Ring et al., 2023; Wikoff, 2023) in a special issue of this journal. The Bayesian methodology, workflow and peer review recommendations were presented and discussed at the beginning of the WHO expert meeting and was thereafter accepted by the panel to determine the 2022 WHO-TEFs accordingly.

6. Bayesian dose response (DR) modeling

A prerequisite in the TEF methodology is the assumption that dose-response curves for all DLC's are parallel and attain the same maximum efficacy. Moreover, this method also assumes that all toxicological and biochemical endpoints for a specific congener have the same REP. In practice, this is more often the exception rather than the rule. Many experimental studies have shown that maximum efficacies are not reached for congeners tested when compared with the reference compound (2,3,7,8-TCDD or PCB 126). In some studies, the test congeners had a greater maximal response compared with the reference compounds. Also, slopes of the dose-response curves are frequently not parallel with that of the reference compound. As a result, previous WHO panels have noted inconsistencies in endpoint specific REPs e.g., between early biological effects such as CYP1A1 induction and more complex toxicological endpoints, such as cholangiocarcinoma of the liver.

To address the above problems in dose-response modeling and employ a consistent REP derivation approach that includes all data sets, a Hill dose-response model was used to estimate REPs and their associated uncertainties using Bayesian estimation (Ring et al., 2023). The Hill model consists of four parameters that describe dose-response relationships. This model is described i.e., by the following formula:

$$y = B + \frac{T - B}{1 + \widehat{10}[(\log_{10} EX50 - \log_{10} x)\ H]}$$

where y is the response, B describes the baseline response, T is the maximal response to the test or reference congeners, x represents either concentration (in vitro) or dose (in vivo), EX*50* is the concentration (C) (for in vitro) or dose (D) (for in vivo) at which 50% of the maximum response is attained, and H represents the Hill slope (the steepness of the curve).

For better understanding, Fig. 1 provides a visual representation of these different parameters and how these affect the shape of the dose-response curve. After normalization to background, the baseline response should be shared by all congeners, including the reference because it reflects the response at zero dose for any chemical. Only the *EX50* could differ between a test congener and the reference chemical with the baseline (B), maximum response (T) and the dose response slope (H) being similar. However, these parameters for DLC's are often not similar between the reference compound and test chemicals. The results of the Bayesian dose-response fit approximate the uncertainty distribution of the different Hill model parameters for each DLC congener.

7. Machine learning dataset quality categories

A qualitative weighting scheme proposed by Wikoff et al. (submitted) identified six attributes that are important criteria for study quality

[3] The development of this framework led by Laurie Haws and Daniele Wikoff and done in collaboration with the 2005 WHO TEF panel members Drs. Linda Birnbaum, Michael DeVito, and Nigel Walker, with additional insight from Drs. William Farland, Martin van den Berg, Michael Denison, Richard Peterson, and Annika Hanberg.

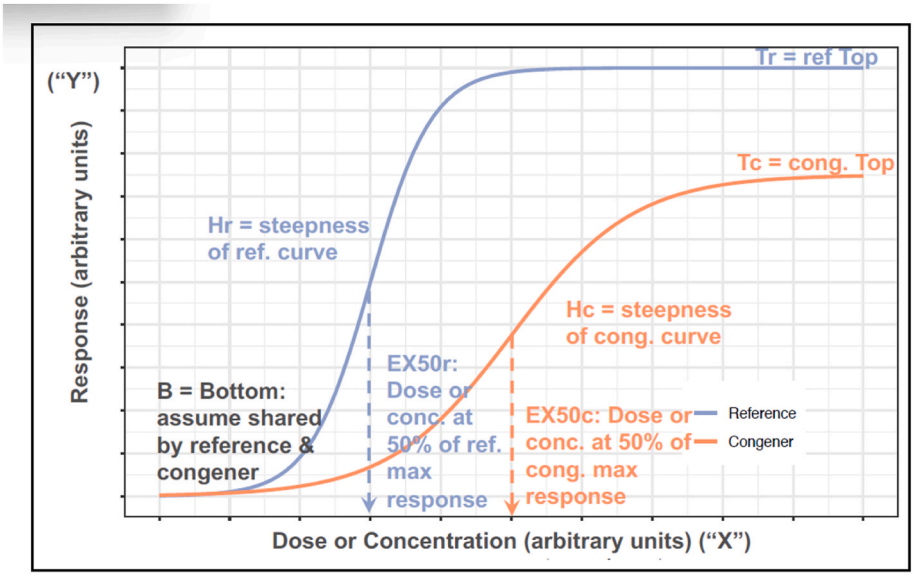

Fig. 1. The parameters used by the Bayesian Hill model to describe the shape of the reference (Blue) and congener (Orange) dose response curves.

and relevance to REP derivation (Wikoff, 2023). Based on expert judgment of these attributes, a panel of experts assigned studies to quality categories from 1 to 5.5, with 1 being the highest score. Studies receiving a score of 5.5 were excluded from the analysis and consisted of QSAR models and studies that did not examine an Ah receptor mediated effect. This set of expert-categorized studies was used as a training set for a machine-learning model to automatically assign quality categories based on study attributes and in fact reproducing the judgment of the expert panel. The output of such an analysis is a probability that a dataset will be assigned to a quality category. Subsequently, the quality category probabilities are used to apply quantitative weighting in a weighted Bayesian Meta-Analysis phase of the workflow. Overall, the

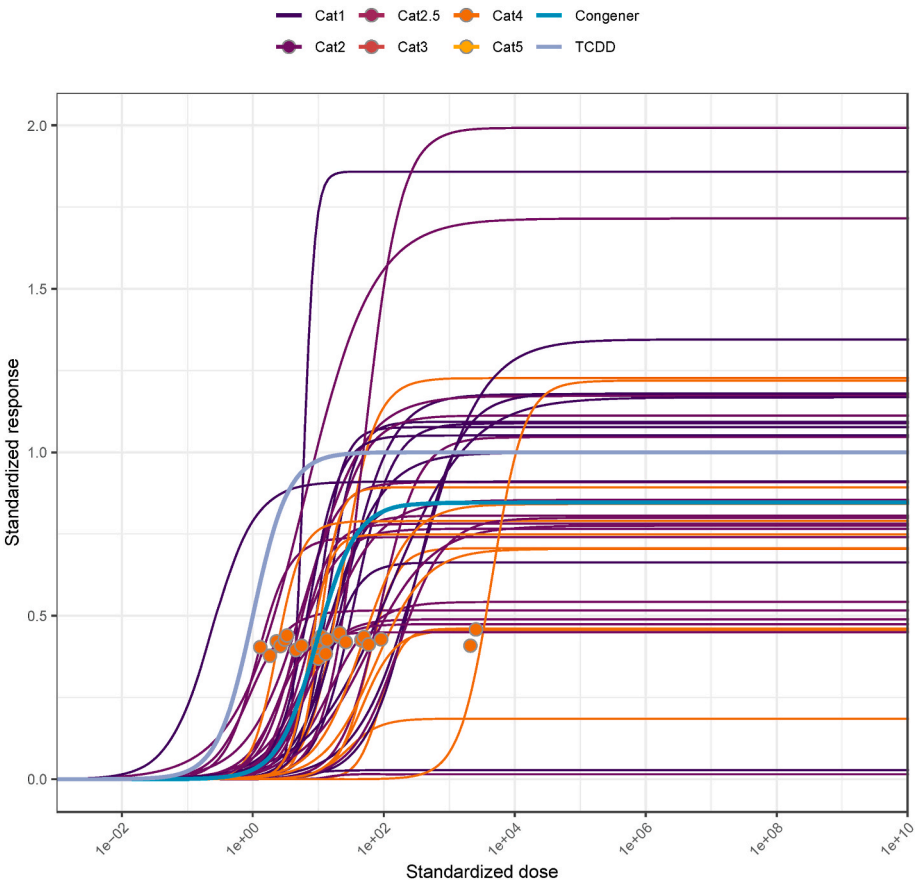

Fig. 2. The above figure shows dose-response curves estimated from individual study data, which were used to estimate individual study REPS. Points represent studies that provide REPS but do not provide dose-response data. Of note is that the maximal response of the test congener often differs from that of TCDD and in some data sets the response to the test congener is greater than TCDD.

results from this approach indicate that there was more certainty in quality categorization for studies assigned a category 1 or 5 compared to categories 2, 3, and 4. Of the six study attributes identified in (Wikoff, 2023) study type (in vivo vs in vitro), REP derivation method, and similarity in pharmacokinetics between the test congener and the reference chemical, contributed the most to the quality categorization.

8. Bayesian Meta-Analysis

Once each data set has gone through dose response modeling, standardization and machine learning quality weighting, a Bayesian meta-analysis was performed by combining all available data for each congener. Included in this analysis were also data sets that provided author-calculated single point estimates in cases where dose response data were not provided. This analysis was performed using weighted and unweighted datasets. This meta-analysis assumes that there is a "true" relationship between the test and reference congener. By integrating all the data from the different sources, a single dose response was then fitted to the data by deriving congener-specific standardized Hill model parameters for individual congeners and reference compounds. Author-derived REP constitutes one individual measurement of a combination of the REP-specific standardized DR parameters, with an unknown (but REP-specific) amount of error (Ring et al., submitted). The result of this meta-analysis has been designated as "Best-Estimate TEF (BE-TEF)" with surrounding uncertainty distributions.

For a given congener, the parameter distributions from the standardized REP specific DR and point estimate REP data were combined with the weight of individual REPs being determined by the quality weightings assigned in the machine learning stage of the workflow. Figs. 2 and 3 present an example of this approach for a single congener. Of note is that the data is standardized so that TCDD has a maximal response of 1 and the test data is standardized to the TCDD dose response curves. As is often the case, the maximum response of a test congener often differs from that of the standardized TCDD maximum as shown in Fig. 2. For author derived point estimate REPs, the Bayesian fits to the individual datasets are presented in Fig. 2 along with a color-coded representation of their weight category. In this approach datasets with higher weights have more influence on the model's prediction of the relationship between the test and reference congener in the weighted analysis. Whereas in an unweighted variation, it integrates each dataset in such a way that no dataset can have more weight than another. Using these datasets and model fits, the output of the Bayesian meta-analysis is in line with the output of the Bayesian DR analysis for each congener (Fig. 3). It includes uncertainty distributions of the standardized Hill parameters that describe the true underlying dose-response relationship for each congener (Fig. 3). Thus Fig. 3 is derived from the data and model fits from Fig. 2. The BE-TEF is derived from the median parameter estimates, represented by the solid blue line in Fig. 3, that results also in a median "Best-Estimate" relationship between the test and reference congener, TCDD.

9. Uncertainties surrounding BE-TEFs

The BE-TEF and its degree of uncertainty in the model parameters (i.e., the size of their respective uncertainty distributions) is characterized by the Credible Interval (CI) surrounding the Best Estimate and is represented using violin plots, an example of which is shown in Fig. 4. In this figure the predicted BE-TEF value is shown as a diamond and represents where the probability density is at its highest (median) of the predicted BE-TEF as indicated by the height of the violin plot. The range of the TEF, represented by the width of the violin tends to be larger where the data is either limited or heterogenous. The violin plot represents the Bayesian estimate (diamond) and the 90% Credible Interval (violin) of BE-TEF values.

In the present analysis, estimates of the credible interval should be viewed as unreliable estimates of the true uncertainty and should not be used for risk management purposes. There are multiple reasons for this. First, the prior distributions over the TEF were based on mathematical plausibility, not toxicological plausibility, and ranged from 1×10^{-20} to 1×10^{6}, and thus, may inflate the uncertainty in data-poor situations. Additionally, the present model does not incorporate potential correlations among REPs for different endpoints or congeners measured by the same laboratory, which will change the uncertainty estimates but not the BE-TEFs estimates. For example, one can still take a mean of multiple correlated observations disregarding the correlation. With enough samples, this mean will accurately estimate the center; however, the uncertainty on the mean estimates will be incorrect because one did not account for the correlation. Despite these limitations, the graphical presentation of these uncertainty estimates (Fig. 5) gives some understanding of the differences in uncertainty between congeners.

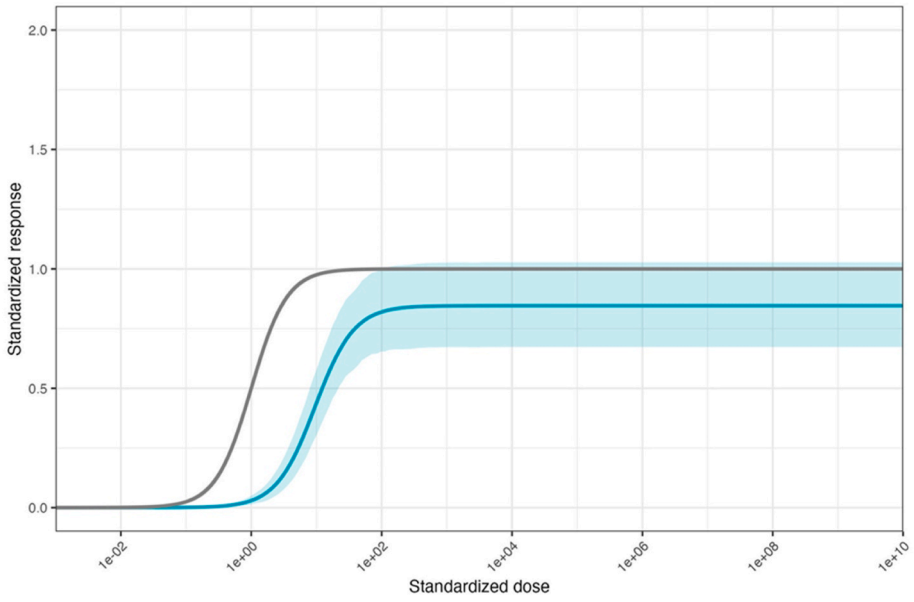

Fig. 3. Modeled reference-congener relationship derived from the Meta-analysis of datasets from Fig. 2. The solid black line represents the TCDD dose response relationship, the blue line represents the best estimate dose response curve for the test congener and the shaded area represents the distribution of the model fits.

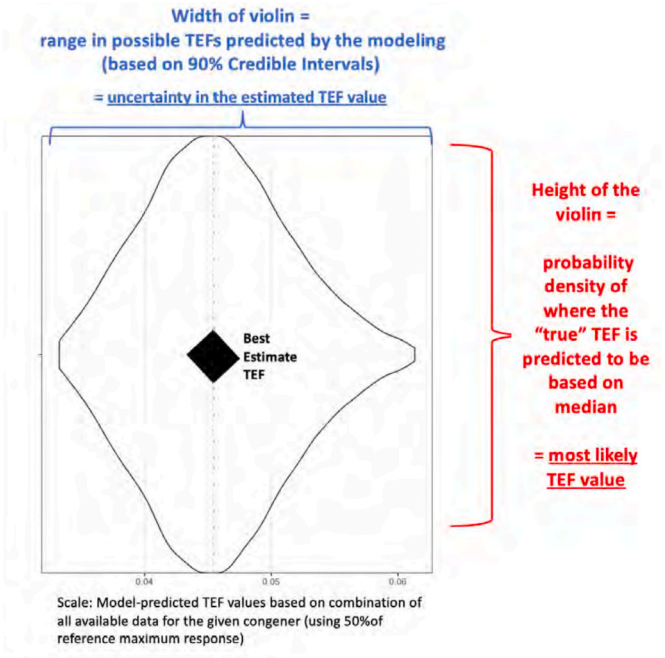

Fig. 4. Violin plot visualizations for interpreting BE-TEF data.

10. Results of the evaluation

10.1. Comparison of 2022 BE-TEFs to 2005 WHO TEFs

Following the above-described machine learning process and Bayesian methodology BE-TEFs were derived for all congeners and the results presented and discussed at the WHO expert meeting of 2022. The outcomes of these calculations were critically evaluated, as well as the relevance of several experimental studies that were (not) included in the Bayesian analysis. This resulted in the acceptance or rejection of some studies with DLC's based on mechanistic or experimental aspects. Several studies evaluating endpoints not clearly mediated through the Ah receptor were excluded from the Bayesian analysis based on recommendations of the panel. These studies examined cell proliferation, cell viability, prostate specific antigen and estrogen and androgen activation and inhibition. Consequently, additional, and renewed calculations via this method were made during and shortly after the meeting.

Following the renewed calculations, the Panel accepted the BE-TEFs as the 2022 WHO-TEFs with the exception for the mono-ortho PCBs. In general, the mono-ortho PCBs tend to have limited data and broad estimates of uncertainty. In addition, the present model excluded studies that demonstrated the mono-ortho PCBs were inactive as dioxin-like chemicals. Given these limitations, the Panel chose to not change the 2005 WHO-TEF values for these congeners. In Table 1 and Fig. 5 the 2022 WHO-TEFs coming out of this expert meeting are presented, including observed uncertainties, and a comparison with the 2005 WHO-TEFs (Van den Berg, Birnbaum et al. 2006). It should be noted that

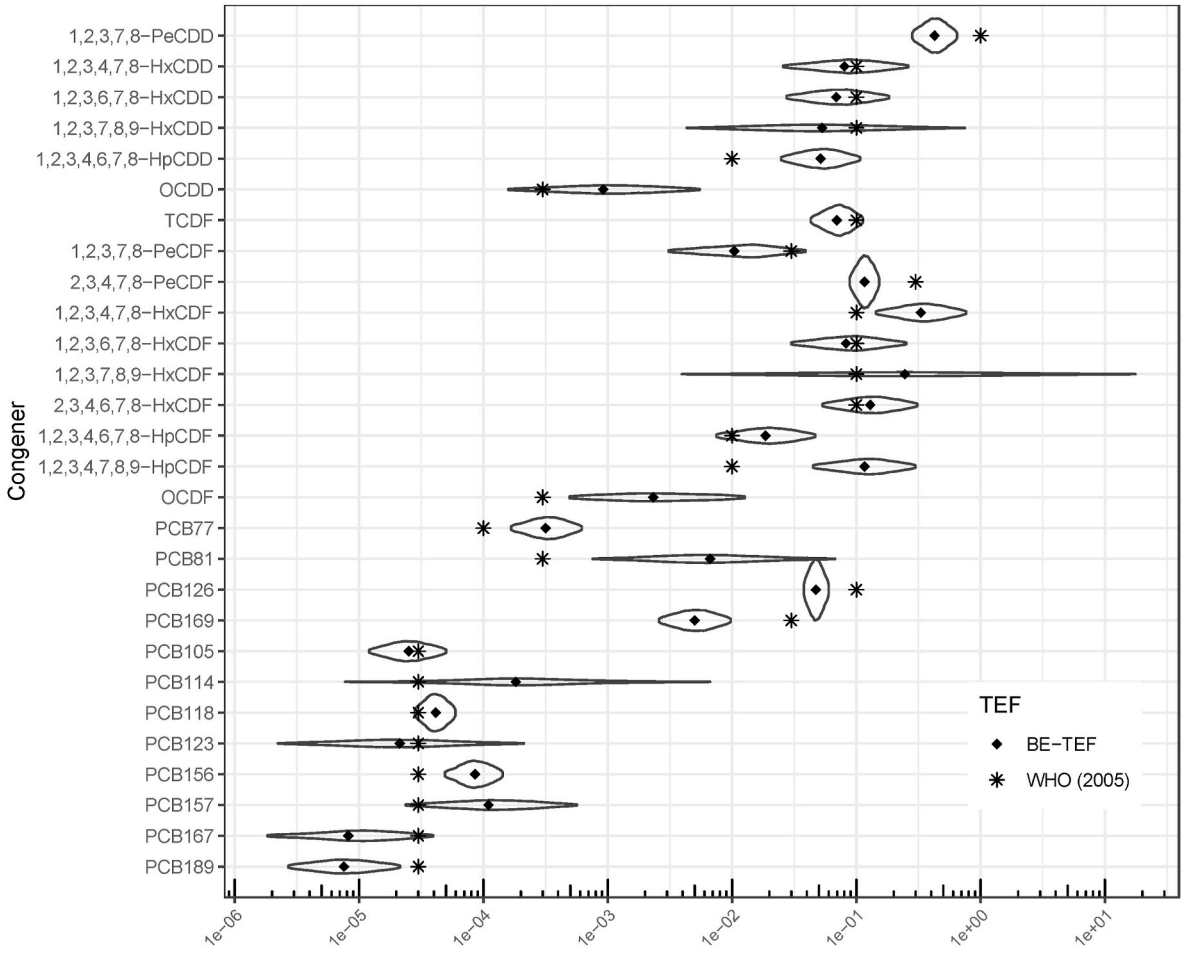

Fig. 5. Best-Estimate TEF estimates (diamonds) and their surrounding model uncertainty distributions (violins), truncated at their respective lower and upper 90% CI. The 2005 WHO TEF (asterisk) is also shown for comparison. The uncertainty of the model is not thought to represent the actual uncertainty, and thus should not be used for risk management purposes. Figure is from Fitch et al. (2023) and is presented with permission.

Table 1
Proposed 2022 WHO-TEFs (BE-TEFs) for dioxin-like compounds and comparison with 2005 WHO-TEFs.

DL Congener	2005 WHO-TEF	BE-TEF 2022	2022 WHO-TEFs	Order of Magnitude Uncertainty
Dioxins				
1,2,3,7,8-PeCDD	1	0.4	**0.4**	½
1,2,3,4,7,8-HxCDD	0.1	0.09	**0.09**	1
1,2,3,6,7,8-HxCDD	0.1	0.07	**0.07**	1
1,2,3,7,8,9-HxCDD	0.1	0.05	**0.05**	2
1,2,3,4,6,7,8-HpCDD	0.01	0.05	**0.05**	½
OCDD	0.0003	0.001	**0.001**	2
Furans				
TCDF	0.1	0.07	**0.07**	½
1,2,3,7,8-PeCDF	0.03	0.01	**0.01**	½
2,3,4,7,8-PeCDF	0.3	0.1	**0.1**	½
1,2,3,4,7,8-HxCDF	0.1	0.3	**0.3**	½
1,2,3,6,7,8-HxCDF	0.1	0.09	**0.09**	**1**
1,2,3,7,8,9-HxCDF	0.1	0.2	**0.2**	2
2,3,4,6,7,8-HxCDF	0.1	0.1	**0.1**	1
1,2,3,4,6,7,8-HpCDF	0.01	0.02	**0.02**	1
1,2,3,4,7,8,9-HpCDF	0.01	0.1	**0.1**	1
OCDF	0.0003	0.002	**0.002**	1
PCBs				
PCB77	0.0001	0.0003	**0.0003**	½
PCB81	0.0003	0.006	**0.006**	2
PCB126	0.1	0.05	**0.05**	1
PCB169	0.03	0.005	**0.005**	½
MONO-ORTHO PCBs				
PCB105	0.00003	0.00003	**0.00003**	3
PCB114	0.00003	0.0002	**0.00003**	3
PCB118	0.00003	0.00004	**0.00003**	3
PCB123	0.00003	0.00002	**0.00003**	3
PCB156	0.00003	0.00009	**0.00003**	3
PCB157	0.00003	0.0001	**0.00003**	3
PCB167	0.00003	0.000009	**0.00003**	3
PCB189	0.00003	0.000008	**0.00003**	3

the newly presented 2022 WHO-TEFs are different from those established in 2005 for many of the congeners (See Table 1). However, for most congeners this difference is less than a ½ log. In part this difference may be attributed to the ½ log 10 rounding that was assigned for the 2005 WHO-TEFs (Van den Berg, Birnbaum et al. 2006). Congeners whose BE-TEFs fall outside of ½ log difference are 1,2,3,4,6,7,8- HpCDD, OCDD, 1,2,3,4,7,8,9-HpCDF, OCDF, PCB81, PCB114, PCB157, PCB167, and PCB169. Preliminary uncertainty estimates around the BE TEFs overlap with the 2005 WHO-TEF except for 1,2,3,7,8-PeCDD, 1,2,3,4,6,7,8-HpCDD, 2,3,4,7,8-PeCDF, 1,2,3,4,7,8-HxCDF, 1,2,3,4,7,8,9-HpCDF, OCDF, PCB77, PCB81, PCB126, PCB169, PCB156, and PCB189 (Table 1). This analysis is presented in detail in Ring et al. (2023) and Fitch et al. (2023). While Table 1 and Fig. 5 present uncertainty ranges for the different congeners, as described previously, they should be used cautiously. Future refinements of the model should include more toxicologically based a priori assumptions and correlations among REPs developed within the same study and laboratory.

10.2. Charge questions to the panel

In view of the number of years between the 2005 and present evaluation of WHO TEF values for DLC's, several charge questions focusing on the science underpinning the TEF methodology were given to the participants of the panel. As the time for the meeting was restricted to five days, it was decided to set up three different breakout groups, each addressing specific questions.

The overall topics of these three subgroups involved.

- *Mechanistic aspects relevant for the TEF concept (I).*
- *The use of* in vitro *studies for TEF derivation (II).*
- *The need for development and use of systemic TEFs (III).*

The three breakout groups reported their findings to the whole expert panel to facilitate a plenary discussion. Tables 2–4 present these questions and the concise answers which were reached by consensus from the whole expert panel. Briefly, the conclusions from these subgroups were as follows.

1. The newly proposed WHO-TEFs have been determined mainly from experimental species (predominantly rodents). The human data included in the TEF derivation was limited to a few in vitro studies examining CYP1A induction. All available human in vivo studies were considered inadequate from an experimental and mechanistic point of view.
2. There are significant differences between rodent and human in vitro REPs, but the present human data are considered inadequate to solely use for risk assessment. These human data are based on biochemical effects, which could be used as markers of elevated AhR-

Table 2
Charge questions to subgroup 1 related to mechanistic aspects relevant for the TEF concept.

Questions	Consensus opinion
Are there human sensitivities/differences that need to be considered when selecting/deriving updated TEF values?	There is clear evidence that the AhR across species has important functions involving normal development and disease processes by regulating cell growth and differentiation in a variety of tissues. There is significant evidence that induction of AhR-related gene expressions determines the degree of toxicity in fish, bird, and mammalian species. Presently our understanding of the biology of AhR activation and the available data, do not provide support to solely apply human in vitro data in selecting or deriving updated TEFs
Is it valid to use human CYP1A1 as the only marker for TEFs related to toxicity in human cells (hepatic, keratinocytes, PBLs)	CYP1A1 cannot be considered as a key element in the toxicity of DLC's. Thus, it may be uncertain to draw quantitative conclusions on the dose-response of critical endpoints and REPs exclusively from data on CYP1A1 induction.
Is there evidence from AhR biology that humans are less sensitive in response for TCDD and PCB126?	There is evidence that for CYP1A1 induction in human primary cells, humans are less sensitive to PCB 126 than rodents. However, as discussed above, the panel does not have confidence that this difference can be directly extrapolated to toxic responses of DLC's.
Is effect on additivity in TEF concept by partial agonist influencing the validity to use it for risk management purposes?	While complex mixtures of DLC's, which also include partial AhR agonists such as mono-ortho PCBs, would be expected to follow an additivity approach, there are limited studies available to provide confirmation. Mixture studies with full AhR agonists, including dioxins, dibenzofurans and non-ortho PCBs have been shown to follow additivity

Table 3
Charge questions to subgroup 2 related to the use of in vitro studies for TEF derivation.

Questions	Consensus Opinion
Are there scientifically sound arguments to use only in vitro rodent data for risk assessment?	No, there are noted differences between rodent in vivo REPS and rodent in vitro REPs. However, the directionality of these differences was inconsistent between congeners and could not be readily explained. Therefore, relying on in vitro only has considerable uncertainty
Are there scientifically sound arguments to use or include in vitro human data for risk assessment?	Human in vitro data may be a good predictor of human vivo responses. It was noted that in vitro potencies are dominated by CYP induction. Comparison of in vivo biochemical and in vivo toxic endpoints (all in rodents) show no indications that these were different. Based on this, it could be concluded that in vivo biochemical responses seem to reflect in vivo toxic responses. However, in the absence of in vivo human evidence to support or reject this for humans, caution is advised for use of in vitro human data on CYP induction alone to derive a TEF.

Table 4
Charge questions to subgroup 3 related to the need for development and use of systemic TEFs.

Questions	Consensus opinion
From a risk assessment point of view is there a need for systemic TEFS?	Systemic TEFs would be useful for biomonitoring and risk assessment based on blood plasma (or extrahepatic) concentrations. Systemic TEFs of congeners may differ from intake TEFs, which warrants future development/ research on systemic TEFs. Relatively small differences between systemic and intake TEFs may be magnified when multiple individual TEFs come together in a systemic TEQ.
Could the present 2022 BE TEFs be used as systemic TEFs in risk assessment?	In the absence of sufficient data to derive systemic TEFs based on blood or other extrahepatic matrices, the 2022 BE TEFs could be used. However, there should be a reference to the uncertainty associated with using these 'intake' TEFs to assess potential health effects based on a systemic concentration.
Is there a need to develop a specific set of systemic TEFs for the interpretation of human biomonitoring data and risk assessment based on plasma concentration (or other extrahepatic matrices)?	The available data shows a difference between intake and systemic TEFs for some congeners, which can be explained by toxicokinetic characteristics. However, there are currently insufficient data to establish systemic TEFs.

mediated activity. However, our present scientific knowledge is not sufficiently robust to use these biochemical activities to link to human toxicity and determine quantitative risks.

3. Systemic TEFs would be useful for biomonitoring and risk assessment based on blood plasma (or extrahepatic) concentrations. Systemic TEFs of congeners may differ from intake TEFs, but there are currently insufficient data to establish systemic TEFs.

It should be noted that by no means do these questions and answers fully cover all scientific aspects related to the TEF concept used for DLC's and the discussion held. However, these charge questions were deemed important for the 2022 WHO re-evaluation of TEFs based on ongoing discussion in the scientific community and amongst risk management experts. In addition, significant time was spent by the expert panel on the use and results of the Bayesian methodology for the 2022 WHO-TEF values.

11. Discussion

Previous WHO meetings used expert judgement in assigning TEFs to DLC's (Ahlborg et al., Van den Berg et al., 1998, 2006). Based on this, a limited number of studies was selected which strongly influenced the derivation of a specific TEF. By doing so, scientific information possibly relevant for human risk assessment of DLC's was not considered. Moreover, the past rounding of a REP value into a TEF value may have insufficiently described the uncertainty around that value. Although this uncertainty was first discussed in the 2005 WHO-TEF reevaluation (van den Berg et al., 2006), it could not be considered during the 2005 expert panel meeting and, as such, was identified as a future improvement to the TEF methodology. Thus, despite the ongoing discussions about differences in species sensitivity, humans included, for DLC's (Black et al., 2012; Larsson et al., 2015; Bock 2017, Shi et al., 2019; Xu et al., 2021), the 2022 WHO-TEF reevaluation considered the Bayesian methodology for deriving the BE-TEFs for PCDDs/PCDFs and PCBs that uses all available relative potency data. Several advantages to this approach were recognized and can be summarized as follows.

- *This methodology utilizes available dose-response data from all studies for the chlorinated DLC's.*
- *By using all these data and the Bayesian techniques, the impact of non-parallelism and differences in maximum efficacy in the data are lessened.*
- *Most, if not all, REPs are consistently derived using the same statistical method.*
- *The Bayesian method results in transparent and objective "most likely estimate" of TEFs that could independently be reproduced, thereby significantly reducing any possible impact of bias in expert judgement.*
- *The use of Bayesian techniques and inferences better characterizes the uncertainty of the model estimated 2022 WHO-TEFs.*

While the panel generally recommended acceptance of the BE-TEF derived with the Bayesian model, exceptions were needed for the *mono-ortho* PCBs. There is very limited experimental data for these congeners, and most are of in vitro origin. In addition, these data for the *mono-ortho* PCBs are very heterogenous. Overall, this decreased the confidence in the BE-TEFs for the *mono-ortho* PCBs by the WHO experts significantly. This variability is consistent with the concept that even among DLC's these compounds are selective AhR modulators (SAhRMs) and the broad width of the violin plots is consistent with species/tissue/ response-dependent selectivity of individual congeners (Safe et al., 2020). As a result, the expert panel decided that the *mono-ortho* PCBs could better be treated together as one class of congeners. Moreover, it was concluded that the limited data used in the Bayesian analysis did not convincingly show a deviation from the 2005 WHO-TEFs. Thus, any TEF change for this class of congeners could not be supported from a scientific point of view. Therefore, it was recommended to retain the 2005 WHO-TEFs for *mono-ortho* PCBs and being similar for this whole class of DLC's.

In vitro human cellular systems have provided information about differences in REPs of non-ortho PCBs with those derived from in vivo and in vitro rodent systems. Several studies using human primary cell systems have indicated that the REP e.g. PCB 126 is significantly lower than those observed in rodent in vivo and in vitro systems. Hiowever, The role of the biochemical changes evaluated in the human in vitro models in the toxicity of TCDD remains uncertain. While the role of CYP1A induction has been proposed as a key event in the mode of action of TCDD rodent carcinogenicity (Budinsky et al., 2014), there is insufficient evidence to supporting a role of CYP1A1 induction in human toxicological responses to TCDD. Specifically, the hepatic tumorigenic response in rodents is significantly different from humans as illustrated

by the fact that liver tumors are not considered a sensitive or relevant endpoint illustrated by many epidemiological studies. Moreover, CYP1A1 is not a major constitutive P450 in the human liver and recently it was shown by Lang et al. (2019) that it's expression is highly variable in human individuals. Thus, using CYP1A1 activity as a biomarker for *human* DLC toxicity appears to be insufficiently supported by the present scientific state of the art. Therefore, the workshop decided to use the same consistent systematic approach for all DLC congeners, including non-ortho PCBs.

As indicated from Table 1, the 2022 WHO-TEFs show a decreasing TEF for many individual DLC's which may have a direct impact on human risk assessment of DLC's. This impact is highest for those congeners that still form a significant part of the total TEQs in e.g., food and feed. In this respect, the changes in TEFs for the non-ortho planar PCBs, most notably for PCB126, as well as those for 1,2,3,7,8-PeCDD and 2,3,4,7,8-PeCDF will contribute the most. It was calculated for several relevant human food products that the total TEQs will approximately decrease by a factor 2 (ca. 50%) using the newly proposed 2022 WHO-TEFs compared with those established in 2005 (see Table 5).

The proposed 2022 WHO-TEFs are the result of new data on the toxicity of DLC's and an improving statistical methodology such as Bayesian analysis**.** The latter aspect clearly provides a more precise indication regarding the uncertainties surrounding the WHO-TEFs values, an aspect that has been regularly brought forward within the scientific community during the last decades. When considering the observed variations in REP values, it should be recognized that there is no group of environmental contaminants that has been studied in so much detail, providing a wealth of experimental data. Inevitably, with so much data available and large differences in experimental designs, the variation surrounding a TEF value becomes more and more significant. It is important to realize that the dynamics in congener specific TEFs are the results of new insights in knowledge about a group of compounds that has been well studied compared with other groups of environmental chemicals.

This new approach to assess relative potencies has several advantages. The development of the weighting scheme prior to the analysis and TEF derivation provides a transparent and systematic assessment of the quality and relevance of individual datasets. The application of the Bayesian dose response modeling and meta-analysis results in an unbiased and transparent TEF value with a quantified uncertainty. Lessons learned from developing the methodology described here for chlorinated DLC TEFs could serve as a model for other groups of chemicals moving forward, such as per- and polyfluorinated substances, polycyclic aromatic compounds, flame retardants, natural toxins, and biocides. Application of this approach to other classes of chemicals is possible but will require modification with particular attention to the weighting scheme. The endpoints used for dioxin-like chemicals would not necessarily apply to other classes of chemicals. In addition, the suitability of the Hill model to endpoints studied for other chemical classes should also be assessed.

Contribution of authors

MdV: were invited WHO independent experts, were involved with writing the first draft of the manuscript. BB: were invited WHO independent experts. MvD: were invited WHO independent experts. KvE: were supporting participants and responsible for the working document use. at the meeting. MF: were invited WHO independent experts, were involved with writing the first draft of the manuscript. EAFG: were supporting participants and responsible for the working document use. at the meeting. LH: were supporting participants and responsible for the working document use. at the meeting. SK: were invited WHO independent experts. REP: were invited WHO independent experts. RH: were invited WHO independent experts. KN: were invited WHO independent experts. KP: coordinated and organized the meeting on behalf of the WHO. CR: were invited WHO independent experts. MR: were invited WHO independent experts. SS: were invited WHO independent experts. DS: were invited WHO independent experts. MW: were invited WHO independent experts. DW: were supporting participants and responsible for the working document use. at the meeting. BZ: were invited WHO independent experts. MvdB: were invited WHO independent experts. chaired this WHO meeting, while MF and MdV served as rapporteurs, were involved with writing the first draft of the manuscript. All authors had the opportunity to comment of the manuscript during two rounds of revisions before submission. All authors were participating in the WHO expert meeting.

Funding body information

I can confirm that WHO has been working with EFSA in order to share the financial burden updating the WHO TEF values for dioxin and dioxin like PCBs. The preparatory work done by KeyToxicology and ToxStrategies including the update of the database was financed by EFSA. The selection of the invited experts was done solely by WHO and the costs for organizing the expert consultation was covered by WHO funds.

WHO has no financial interests in the outcome of the newly updated 2022 WHO TEF values for dioxin and dioxin-like PCBs. The update was initiated by a request from the Joint FAO/WHO Codex Alimentarius Commission on Contaminants in Food to JECFA and WHO to prepare an updated dietary safety assessment of dioxin and dioxin like PCB's in food. Before an updated JECFA assessment on dioxin etc. Can be initiated it was found critical to first update the previous WHO TEF values.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

WHO has been working with EFSA in order to share the financial burden to update the TEF values for dioxin etc. The preparatory work done by KeyToxicology (KvE, EAFG) and ToxStrategies (LH, DW) including the update of the database was financed by EFSA. The selection of the invited experts was done solely by WHO and the costs for organizing the expert consultation was done by WHO funds. None of the WHO invited experts have indicated any conflicting interests in their WHO declarations of interests which were evaluated prior to the meeting.

Table 5
Changes in total concentration of TEQs (pg/g lipid) for human milk and fish products when the 2022 WHO-TEFs are used compared to the 2005 WHO-TEFs.

TEQs	(Houlihan et al., 2022) Human milk	(Ferreira and Moreira Mde 2015) Human milk	(Perello et al., 2015) Sardines	(Zacs et al., 2013) Eel	(Hasegawa et al., 2007) Fish oils
Dioxins/Furans	3.1 pg/g (2005)	15.57 pg/g (2005)	0.25 pg/g (2005)	1.95 pg/g (2005)	2.1 pg/g (2005)
	2.03 pg/g (2022)	9.14 pg/g (2022)	0.14 pg/g (2022)	1.14 pg/g (2022)	1.35 pg/g (2022)
Dioxin-like	1.55 pg/g (2005)	6.96 pg/g (2005)	1.99 pg/g (2005)	4.75 pg/g (2005)	8.8 pg/g (2005)
PCBs	0.78 pg/g (2022)	3.89 pg/g (2022)	0.99 pg/g (2022)	2.43 pg/g (2022)	4.504 pg/g (2022)
Total 2005	4.65 pg/g	22.33 pg/g	2.239 pg/g	6.7 pg/g	10.9 pg/g
Total 2022	2.81 pg/g	13.03 pg/g	1.134 pg/g	3.57 pg/g	5.85 pg/g
% Decrease	**40**	**42**	**50**	**47**	**46**

Data availability

Data will be made available on request.

References

Ahlborg, U.G., Hanberg, A., 1994. Toxic equivalency factors for dioxin-like PCBs. Environ. Sci. Pollut. Res. Int. 1, 67–68.

Beischlag, T.V., et al., 2008. The aryl hydrocarbon receptor complex and the control of gene expression. Crit. Rev. Eukaryot. Gene Expr. 18, 207–250.

Birnbaum, L.S., 1994. The mechanism of dioxin toxicity: relationship to risk assessment. Environ. Health Perspect. 102 (Suppl. 9), 157–167.

Birnbaum, L.S., Tuomisto, J., 2000. Non-carcinogenic effects of TCDD in animals. Food Addit. Contam. 17, 275–288.

Black, M.B., et al., 2012. Cross-species comparisons of transcriptomic alterations in human and rat primary hepatocytes exposed to 2,3,7,8-tetrachlorodibenzo-p-dioxin. Toxicol. Sci. 127, 199–215.

Bock, K.W., 2017. Human and rodent aryl hydrocarbon receptor (AHR): from mediator of dioxin toxicity to physiologic AHR functions and therapeutic options. Biol. Chem. 398, 455–464.

Budinsky, R.A., et al., 2014. Mode of action and dose-response framework analysis for receptor-mediated toxicity: the aryl hydrocarbon receptor as a case study. Crit. Rev. Toxicol. 44, 83–119.

Denison, M.S., Faber, S.C., 2017. And now for something completely different: diversity in ligand-dependent activation of Ah receptor responses. Curr Opin Toxicol 2, 124–131.

Denison, M.S., et al., 2011. Exactly the same but different: promiscuity and diversity in the molecular mechanisms of action of the aryl hydrocarbon (dioxin) receptor. Toxicol. Sci. 124, 1–22.

Ferreira, A.P., Moreira Mde, F., 2015. Dioxins and furans in breast milk: a case study of mothers from southern Rio de Janeiro, Brazil. Cad. Saúde Pública 31, 1107–1111.

Fitch, S., A, B., et al., 2023. Systematic update to the mammalian relative potency estimate database and development of best estimate toxic equivalency factors for dioxin-like compounds. Regul. Toxicol. Pharmacol. Accepted and press.

Hahn, M.E., et al., 2017. Diversity as opportunity: insights from 600 million years of AHR evolution. Curr Opin Toxicol 2, 58–71.

Hankinson, O., 1995. The aryl hydrocarbon receptor complex. Annu. Rev. Pharmacol. Toxicol. 35, 307–340.

Hasegawa, J., et al., 2007. Determination of PCDD/Fs and dioxin-like PCBs in fish oils for feed ingredients by congener-specific chemical analysis and CALUX bioassay. Chemosphere 69, 1188–1194.

Haws, L.C., et al., 2006. Development of a refined database of mammalian relative potency estimates for dioxin-like compounds. Toxicol. Sci. 89, 4–30.

Hays, S.M., Aylward, L.L., 2003. Dioxin risks in perspective: past, present, and future. Regul. Toxicol. Pharmacol. 37, 202–217.

Houlihan, M., et al., 2022. Concentrations of polychlorinated dibenzo-p-dioxins (PCDDs), polychlorinated dibenzofurans (PCDFs) and polychlorinated biphenyls (PCBs) in human milk from Ireland: temporal trends and implications for nursing infant exposure. Journal of Environmental Exposure Assessment 1, 2.

Lang, D., et al., 2019. Highly variable expression of CYP1A1 in human liver and impact on pharmacokinetics of Riociguat and Granisetron in humans. Chem. Res. Toxicol. 32, 1115–1122.

Larsson, M., et al., 2015. Consensus toxicity factors for polychlorinated dibenzo-p-dioxins, dibenzofurans, and biphenyls combining in silico models and extensive in vitro screening of AhR-mediated effects in human and rodent cells. Chem. Res. Toxicol. 28, 641–650.

Loomis, D., et al., 2018. Identifying occupational carcinogens: an update from the IARC Monographs. Occup. Environ. Med. 75, 593–603.

Muzembo, B.A., et al., 2019. Dioxins levels in human blood after implementation of measures against dioxin exposure in Japan. Environ. Health Prev. Med. 24, 6.

Perello, G., et al., 2015. Human exposure to PCDD/Fs and PCBs through consumption of fish and seafood in Catalonia (Spain): temporal trend. Food Chem. Toxicol. 81, 28–33.

Ring, C., et al., 2023. A multi-tiered hierarchical Bayesian approach to derive toxic equivalency factors for dioxin-like compounds. Regul. Toxicol. Pharmacol. 143, 105464.

Safe, S., 1990. Polychlorinated biphenyls (PCBs), dibenzo-p-dioxins (PCDDs), dibenzofurans (PCDFs), and related compounds: environmental and mechanistic considerations which support the development of toxic equivalency factors (TEFs). Crit. Rev. Toxicol. 21, 51–88.

Safe, S.H., 1994. Polychlorinated biphenyls (PCBs): environmental impact, biochemical and toxic responses, and implications for risk assessment. Crit. Rev. Toxicol. 24, 87–149.

Safe, S., et al., 2020. Aryl hydrocarbon receptor (AHR) ligands as selective AHR modulators (SAhRMs). Int. J. Mol. Sci. 21.

Shi, H., et al., 2019. Concentration dependence of human and mouse aryl hydrocarbon receptor responsiveness to polychlorinated biphenyl exposures: implications for aroclor mixtures. Xenobiotica 49, 1414–1422.

USEPA, 2010. Recommended Toxicity Equivalence Factors (TEFs) for Human Health Risk Assessments of 2,3,7,8-Tetrachlorodibenzo-P-Dioxin and Dioxin-like Compounds. USEPA, Washington D.C.

USEPA, 2012. EPA's Reanalysis of Key Issues Related to Dioxin Toxicity and Response to NAS Comments, vol. 1. USEPA, Washington D.C,.

van den Berg, M., et al., 1989. Selective retention of toxic polychlorinated dibenzo-p-dioxins and dibenzofurans in the liver of the rat after intravenous administration of a mixture. Toxicology 55, 173–182.

Van den Berg, M., et al., 1998. Toxic equivalency factors (TEFs) for PCBs, PCDDs, PCDFs for humans and wildlife. Environ. Health Perspect. 106, 775–792.

Van den Berg, M., et al., 2006. The 2005 World Health Organization reevaluation of human and Mammalian toxic equivalency factors for dioxins and dioxin-like compounds. Toxicol. Sci. 93, 223–241.

Wikoff, D, Ring, C, DeVito, M, Walker, N, Birnbaum, L, Haws, L, 2023 Oct 21. Development and application of a systematic and quantitative weighting framework to evaluate the quality and relevance of relative potency estimates for dioxin-like compounds (DLCs) for human health risk assessment. Regul. Toxicol. Pharmacol. 145, 105500. https://doi.org/10.1016/j.yrtph.2023.105500.

Xu, X., et al., 2021. Species-specific differences in aryl hydrocarbon receptor responses: how and Why? Int. J. Mol. Sci. 22.

Zacs, D., et al., 2013. Content of polychlorinated dibenzo-p-dioxins, dibenzofurans and dioxin-like polychlorinated biphenyls in fish from Latvian lakes. Chemosphere 91, 179–186.