Food and Chemical Toxicology 182 (2023) 114182

Contents lists available at ScienceDirect

Food and Chemical Toxicology

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

Update of the Cancer Potency Database (CPDB) to enable derivations of Thresholds of Toxicological Concern (TTC) for cancer potency

Chihae Yang [a,b,c], James F. Rathman [a,b,c], J. Vinicius Ribeiro [b], Monika Batke [d], Sylvia E. Escher [d], James W. Firman [e], Bryan Hobocienski [b], Rupert Kellner [d], Aleksandra Mostrag [b], Katarzyna R. Przybylak [e], Mark T.D. Cronin [e,*]

[a] Molecular Networks GmbH, Nuremberg, Germany
[b] Altamira LLC, Columbus, OH, USA
[c] The Ohio State University, Columbus, OH, USA
[d] Fraunhofer ITEM, Hanover, Germany
[e] Liverpool John Moores University, Liverpool, UK

ARTICLE INFO

Handling Editor: Dr. Bryan Delaney

Keywords:
Cancer potency database
Threshold of toxicological concern
Point of departure
Benchmark dose

ABSTRACT

The purpose of this study was to update the existing Cancer Potency Database (CPDB) in order to support the development of a dataset of compounds, with associated points of departure (PoDs), to enable a review and update of currently applied values for the Threshold of Toxicological Concern (TTC) for cancer endpoints. This update of the current CPDB, last reviewed in 2012, includes the addition of new data (44 compounds and 158 studies leading to additional 359 dose-response curves). Strict inclusion criteria were established and applied to select compounds and studies with relevant cancer potency data. PoDs were calculated from dose-response modeling, including the benchmark dose (BMD) and the lower 90% confidence limits (BMDL) at a specified benchmark response (BMR) of 10%. The updated full CPDB database resulted in a total of 421 chemicals which had dose-response data that could be used to calculate PoDs. This candidate dataset for cancer TTC is provided in a transparent and adaptable format for further analysis of TTC to derive cancer potency thresholds.

1. Introduction

The Threshold of Toxicological Concern (TTC) is a well-established risk assessment paradigm whereby low exposure to a substance can be deemed to have a high probability of being safe (Embry et al., 2014; Kroes et al., 2004). To apply TTC to a compound, such as an impurity or ingredient, the exposure must be known for the compound to be classified. TTC values are commonly applied to non-cancer and potentially cancer inducing datasets. The context and derivation of the so-called cancer TTC value is well described by Boobis et al. (2017). Briefly, the values were derived, in part at least, from a probabilistic analysis (Rulis, 1987) of the Carcinogenicity Potency Database (CPDB) (Gold et al., 1984, 1989, 1995).

The CPDB was developed between 1980 and 2004 under the direction of Dr Lois Gold, leading the Carcinogenic Potency Project at the University of California, Berkeley, and the E.O. Lawrence Berkeley National Laboratory. It provides quantitative and qualitative results of carcinogenicity studies in rats, mice, hamsters, dogs, and non-human primates for 1547 compounds. CPDB data were compiled from the literature (over 1370 publications published between 1937 and 2003) and National Cancer Institute/National Toxicology Program (NCI/NTP) reports (over 480 reports published between 1976 and 2003) – this is termed the "CPDB 1995" in this paper. For each experiment recorded, the CPDB includes: (i) study design details: species, strain, sex, route of administration, dosing and experiment duration, average daily doses in mg/kg bw/day; (ii) study outcome: target organs, tumor types, and tumor incidences; (iii) carcinogenic potency (Tumorigenic Dose 50 (TD50)) along with its statistical significance; (iv) shape of the dose-response relationship; (v) authors' conclusion regarding carcinogenicity; (vi) detailed citation (literature or NTP report). The in-depth description of methods applied during the development of the CPDB (study inclusion criteria, tissue/tumor types inclusion criteria, standardisation of average daily doses, TD50 estimation and analysis) is available for download (NIH CPDB, 2023).

With regard to developing TTC values, general methods for the

* Corresponding author.
E-mail address: m.t.cronin@ljmu.ac.uk (M.T.D. Cronin).

https://doi.org/10.1016/j.fct.2023.114182
Received 17 October 2023; Received in revised form 3 November 2023; Accepted 3 November 2023
Available online 10 November 2023

Abbreviations:

AIC	Akaike Information Criterion
BMD	Benchmark Dose
BMDL	Lower Benchmark Dose
BMDU	Upper Benchmark Dose
BMR	Benchmark Response
CAS RN	Chemical Abstract Service Registry Number
COC	Cohort of Concern
CMS ID	COSMOS Structure Registry Number
CPDB	Cancer Potency Database
d	dose
ECHA	European Chemical Agency
EFSA	European Food Safety Authority
GLP	Good Laboratory Practice;
IOM	Inorganic Organometallics and Metal complexes
ITEM	Institute for Toxicology and Experimental Medicine;
MTD	Minimum Toxicological Dose
NCI	National Cancer Institute
NLL	Negative Log-Likelihood
NTP	National Toxicology Program
P_d	Proportion of tumor-bearing animals
PoD(s)	Point(s) of Departure
QC	Quality Control
TBA	Tumor Bearing Animals
TD50	Tumorigenic Dose 50
TG	Test Guideline;
TOR	Threshold of Regulation
TTC	Threshold of Toxicological Concern
US FDA	United States Food and Drug Administration

development of the dataset to support the TTC approach have been described previously (Boobis et al., 2017). Briefly, the TTC offers a pragmatic solution for assessing the safety of low-exposure food-contact compounds, fragrances and flavorings, and is also being considered for cosmetics and other types of compounds (Boobis et al., 2017; EFSA, 2016; Yang et al., 2017). TTC was developed based on the Threshold of Regulation (TOR) policy adopted at the United States Food and Drug Administration (US FDA) in 1995 after the evaluation of the range of carcinogenic potencies of 477 compounds included in the CPDB, with the goal being to identify the level of negligible risk across all toxicological endpoints, including carcinogenicity (US FDA, 1995). In the publication by Cheeseman et al. (1999), the minimum oral TD50 values were selected from studies giving certain tumor incidences with p-values equal to or lower than 0.05. The extended CPDB was analyzed in terms of structural alerts and Ames assay data to correlate results with the initial dataset upon which TOR was based, and to identify different potencies between structural classes and the mutagenic potential of chemicals. This analysis demonstrated that carcinogens negative in the Ames test are less potent than those that are positive.

Subsequent work in the development of the data compilation to support TTC was undertaken by Aungst et al. of the USFDA (Aungst et al., 2012). This analysis was based on the most recent update of the CPDB in 2004 (Gold et al., 2005). After curation of the 2004-CPDB database (termed "CPDB 2004"), a TTC dataset was subsequently identified and curated, here referred to as "CPDB 2012" (available for download from https://mn-am.com/demos-services/). Only those CPDB studies compliant with US FDA Redbook 2000 (US FDA, 2000) criteria were considered, namely: Good Laboratory Practice (GLP) studies with relevant protocols, appropriate sample size (>40), dose levels (single dose studies were excluded; preferred Minimum Toxicological Dose (MTD) achieved), and acceptable duration (18 months for mice, 24 months for rats, shorter exposures were included if results were statistically positive). The lowest TD50 values from oral studies with significant tumorigenic effects (p-value $\leq$0.05) were selected. CPDB studies reporting findings of tumor bearing animals (TBA) or mixed-site or mixed-tumor findings, e.g., codes MXA and MXB, as well as negative and equivocal studies, were excluded from the TTC analysis. The resultant CPDB 2012 contained 395 positive compounds selected by the above inclusion criteria.

For the purposes of this paper, it is important to note that Gold et al. (1995) estimated the median toxic dose, TD50, for each dose-response curve (tumor incidence data) reported in the CPDB to provide a standardized quantitative measure for comparisons and analyses of carcinogenesis observations. This numeric descriptor for cancer potency is required to derive the TTC threshold. There have been some concerns on the use of existing TD50 values. For example, Gold et al. (1995) report TD50 values even when a study included data for only one dose; they also report TD50 values when no tumors were observed in the experiment.

In their review and critical assessment of TTC values for compounds that are genotoxic and/or carcinogenic, Boobis et al. (2017) reiterated the conclusions of the European Food Safety Authority (EFSA) (EFSA, 2016), namely, that expanding the cancer TTC dataset would enhance the range of chemical structures. In addition, it would allow for other determinations of points of departure (PoDs) to be applied. This manuscript reports part of the outcome of a project to undertake this task. As such, the aim of this investigation was to update the historic CPDB into a useable resource for TTC, specifically to:

- Include new data in the CPDB
- Define and utilise transparent quality control criteria for the inclusion of data into the database
- Calculate PoDs

From the outset, the purpose of this study was to develop the CPDB such that the data could be used to derive a dataset to support TTC values. A TTC analysis for compounds with a non-genotoxic mode of action on one of the interim datasets has been published elsewhere (Batke et al., 2021) and was the subject of an international workshop (Escher et al., 2023). This current study delivers an updated portion of the CPDB database as well as a new dataset from which a Cancer TTC dataset can be established.

2. Methods

2.1. Identification of new data to update the CPDB including curation of the existing CPDB

The development of the updated database of cancer potency values used as its basis the final 2004-CPDB update "CPDB 2004" (Gold et al., 2005) and the subsequent update by Aungst et al. (2012) (termed "CPDB 2012"). The initial dataset and further explanatory information is available from: https://mn-am.com/demos-services/. The dataset was updated by adding recent studies of genotoxic and non-genotoxic carcinogen data for chemicals from publicly available sources. The study inclusion criteria followed the CPDB database standard (Gold et al., 1984), and the more recent recommendations published by Boobis et al. (2017). The focus was put on *in vivo* studies typically used in regulatory risk assessment, e.g., oral studies involving rodents, dogs, or monkeys. The data were harvested from the sources listed in Table 1. In brief, these were: US NTP (https://ntp.niehs.nih.gov/), Fraunhofer Institute for Toxicology and Experimental Medicine (ITEM) RepDose database (http://fraunhofer-repdose.de/), EFSA (http://www.efsa.europa.eu/) and European Chemical Agency (ECHA) (https://echa.

Table 1
Sources of compounds in the 2004 vs 2023 CPDB databases including multiple studies for individual compounds.

CPDB Database	Literature	NTP	EFSA	ECHA	RepDose
CPDB 2004	1250	443	0	0	0
CPDB 2023 (in this paper)	1	26	12	1	22
Total	1251	469	12	1	22

europa.eu/).

2.2. Quality control of chemical structures and associated carcinogenicity studies

The data quality control (QC) process for the creation of the updated CPDB 2023 included chemoinformatics curation of chemical compounds as well as reviews of toxicity studies. This ensured whether the correct chemical information was associated with the cancer potency data including structures and identifiers. Additional inclusion criteria were enforced to develop the database from which the dataset for TTC was extracted. To this end, test substances in the updated database were constrained to chemical compounds having feasible structure representations. Mixtures or substances with complex compositions were included only when structures could be represented to allow assignment to a chemical class.

The compounds in the database were verified to ensure the correctness and accuracy of how the tested forms (e.g. presence of salts, the neutral forms, etc.) were represented. Compounds were registered in the COSMOS registry system with CMS identifiers (CMS ID) (Yang et al., 2021). The compound records include common identifiers, e.g., Chemical Abstract Service Registry Numbers (CAS RNs), and chemical names from the data sources. The chemical structure representations are available as SMILES strings, INChI codes and keys in the database, with additional elaboration of representative structures in the case of mixtures. The chemical records for the Cancer TTC candidate dataset underwent an additional curation process to ensure their correctness. Inorganic, organometallics, and metal complexes (IOM) were labeled as such. Additionally, all structures were checked using the CORINA Symphony software tool (CORINA Symphony, 2023) to detect potential issues with bond query features, including: incorrect number of free electrons on atoms; incorrectly specified fixed valence or radical state; S-Group logic was used to reflect mixtures, polymers and coordination compounds; hypervalent atoms; invalid charges; pentavalent nitro-groups; multiple fragments; charged heavy or carbon atoms. The validity/invalidity of the reported structures was inspected manually. Chemical identifiers for the data set are provided in Supplementary Files Tables S1 and S2 (Tabs "S1.1" and "S2.1").

The updated CPDB2023 captured the information on the guideline and GLP-compliance as well as on the data record reliability as provided in the data sources. The QC part of the database captures experts' opinions and results of the conducted study reviews, comments on the numeric endpoint values, and recommendations regarding the relevance of the studies for the project.

2.3. Definition of inclusion criteria and application to the CPDB

Criteria were established for selecting data from the updated CPDB 2023 to be used in developing the TTC dataset. The effect-level descriptions in Table 2 were applied to exclude data when calculating the numeric cancer potency measures, the PoDs described in Section 2.4.

With regard to human relevance, in order to avoid losing data, all studies were included even if the study results may indicate that the results are not human relevant. The exception was for the well-known α2μ-globulin nephropathy related cases (Goyak et al., 2022) which were removed from the data set. Opinions with tumor indications included "clear evidence (NTP)", "positive", "some evidence (NTP)", and "associated" were included.

Table 2
Conditions for filtering of effects and exclusion of cancer potency data from the updated CPDB 2023 to allow for the calculation of PoDs.

Fields	Records excluded if
QC notes	the QC note states that no dose level or incidence information was available
Route	the dose was delivered via routes other than oral
Opinion (on the effects)	the tumor effect was tagged as "negative" or "no evidence - NTP"
SITE (3-letter code)	the site code was "mix", "mxb", or "tba (all tumor bearing animal)"
TYPE (3-letter code)	the tumor type code was "mix", "mtm", "MXA (mxa)", or "MXB (mxb)"
Comments	the comments included "age related tumors", "no evidence compared to historical control data", or "not suitable for cancer evaluation"

Overall, these inclusion criteria allowed for the addition of 44 compounds providing 158 studies and 359 dose-response curves) in the updated CPDB 2023. The final counts in the updated CPDB 2023 were 1591 compounds, 6227 studies, and 23,363 dose-response curves. More than 30% of the data of the whole database originated from NTP studies, whereas 60% (26 out of 44 compounds) of the updates were from NTP.

2.4. Calculation of PoDs for the cancer potency data

The cancer potency data that met the inclusion criteria defined in Section 2.3 were subject to analysis described in this section to calculate PoDs. The TTC analysis of the PoDs was not undertaken in this study. It is intended that the PoDs may be subject to TTC analysis elsewhere.

2.4.1. Qualifying study data for benchmark dose modeling

The benchmark response (BMR) is the specified effect level at which the benchmark dose (BMD) will be estimated. For example, specifying a BMR of 10% (BMR = 0.10) indicates we wish to estimate the dose at which 10% of the observations would be expected to have a positive response. BMR is thus a specified value of the probability of a positive response of interest and BMD the dose at which this will occur, hence BMDx (at a specified BMR of x%), should be greater than zero by definition. The goal of the analyses in this study was to report BMD and BMDL results only for studies that meet the relevant criteria and only when the experimental dose-response data indicate clear and unambiguous tumor findings. To achieve this goal, study qualification was based on a clearly-defined set of criteria which were applied in a systematic, reproducible and completely transparent manner. BMDx values are thus reported only when the following criteria were met:

1) The study satisfied the rigorous inclusion criteria that were systematically applied to the full database to extract qualified records. Calculation of a BMDx value from a given set of dose-response curve was not considered if the study was not qualified. These criteria are described in detail in Section 2.3.
2) The experimental data included results at two or more dose levels. Single-dose studies were excluded.
3) Contingency table analysis indicated a significant relationship between dose and tumor counts. The relationship is deemed significant if the p-value for this test is < 0.05. It is important to note that contingency table analysis does not assume that the dose-response data are described by any particular mathematical model. This is, therefore, a model-independent test of the suitability of the dose-response data for subsequent modeling. This criterion screens out studies in which the proportion of animals with tumors does not vary significantly with dose. One such situation in which this may occur is when a significant number of animals with tumors are observed in the control group.

4) The experimental data (dose (d) vs. proportion of tumor-bearing animals (P_d)) were fit using logistic regression with a logit link function:

$$\ln\left(\frac{P_d}{1-P_d}\right) = a + bd \quad (1)$$

The Akaike Information Criterion (AIC) (deLeeuw, 1992) was used to evaluate the quality of the model. AIC is typically used to compare different models applied to the same dataset; the "best" model has the smallest AIC. According to EFSA guidelines (Hardy et al., 2017), "statistical evidence of a dose-related trend" can be claimed if the AIC for a given model is at least 2 less than the AIC for the corresponding null model.

In this case, the null hypothesis is that there is no relationship between d and P_d, so the null model corresponding to the model in eq (1) is:

$$\ln\left(\frac{P_d}{1-P_d}\right) = a \quad (2)$$

To be compliant with EFSA guidance (Hardy et al., 2017), BMDx values were reported only for datasets that, when modeled as described above, give (AIC_{null} – AIC) > 2.

5) Given estimates of the model parameters a and b obtained by the fitting the experimental data to the model in eq (1), BMDx values were calculated at three levels (x = 10%, 25%, 50%):

$$BMD_X = \frac{1}{b}\left(\ln\left(\frac{X}{100-X}\right) - a\right) \quad (3)$$

Studies meeting all other criteria were deemed qualified only if 0 < BMD10 < BMD25 < BMD50. In other words, all BMDx values must be positive and BMDx must increase monotonically with x. This criterion excludes unusual cases in which the proportion of tumor-bearing animals decreases with increasing dose, or the proportion of tumor-bearing animals in the control group is so high that BMDx would be negative (e.g., if more than 10% of animals in the control group are observed to have tumors, then the calculated BMD_{10} would be negative).

Although the logistic model with logit link (eq (1)) is used as described above (criteria 4 and 5) for study qualification, we did not assume this is the best model for fitting the dose-response data. Once qualified studies are identified, benchmark dose modeling was performed using PROAST (RIVM, 2023), which considers multiple models. This is described below.

2.4.2. Benchmark dose modeling

Given a benchmark response BMR of interest (e.g., 10%), the objective of benchmark dose modeling is to estimate values for the corresponding benchmark dose (BMD) and lower and upper confidence limits (BMDL and BMDU).

Benchmark dose modeling was performed using PROAST v70.3 from RIVM (2023) available in RStudio v1.4.1717 (https://posit.co/products/open-source/rstudio/). Modeling was performed as a batch job after setting the PROAST configuration settings and options, and applying study selection criteria systematically. The specific configuration settings used in PROAST modeling are listed in Table 3. The selection of dose response curves followed the rules defined earlier in Table 2 (section 2.3) for study inclusion.

BMD values were calculated with a predefined BMR of 10%. For PROAST calculations, this was specified as the additional risk, so the expected dose at which observed tumor incidence rate is 10% higher than the control (no treatment) group. Only studies meeting the POD inclusion criteria, specified in Section 2.1, and meeting the study qualifying criteria (Section 2.4.1) were submitted for modeling. The names used by PROAST for the nine models considered for each dose-response dataset are: null, full, two.stage, log.logist, Weibull, log.prob, gamma, LVM: Expon m3-, LVM: Hill m3-. Null and full models are reference models whilst the other seven models were averaged per input data. Model averaging was applied using the default settings, meaning that a prior selection of models was not undertaken. PROAST weights the individual models used in model averaging, e.g. models without a significant trend according to the AIC that was used to evaluate the quality of the model. As described before, AIC is a simple measure that takes both NLL (negative log-likelihood) and number of parameters into account:

$$AIC = 2(k + NLL) \quad (4)$$

where k is the number of independent parameters in the model. Models not meeting this criterion are assigned a weight of zero and thus do not contribute to model averaging. Although PROAST can consider covariate variables, none were included in this study. The benchmark dose modeling results are reported as the BMD10 and the associated two-sided 90% confidence interval, namely BMDL10 (lower limit of BMD10), BMDU10 (upper limit of BMD10) for each evaluated dose group.

Table 3
Configuration settings and pre-filtering algorithms for the benchmark dose modeling.

Pre-screening algorithms or criteria	Selection based on
Pre-screening for eq (1) and meeting AIC criteria (EFSA guidance)	Logit model (python)
Condition: 0 < BMD10 < BMD25 < BMD 50	Logit model (python)
Observations: inclusion/exclusion criteria	Table 2
Observations: removal of single dose studies	dose-response data
PROAST Option (ver 70.3/RStudio v1.4.1717) Questions	**Selected option**
What type of response data do you want to consider?	quantal
Do you want to fit a set of models, or choose a single model?	set of models
Which variable do you want to consider as independent variable? (e.g. dose, age	dose
Which response(s) you want to analyze by set of models?	positive.counts
Enter column with the associated sample sizes.	total.counts
Give number of factor serving as potential covariate (e.g., sex)- type 0 if none	0 (none)
What type of Benchmark response do you want to consider? Type 0 if you do not need confidence intervals (CIs).	Additional risk, i.e. P [BMD] – P[0]
Give value for the BMR, in terms of additional risk.	0.1
Do you want to calculate the BMD confidence interval by model averaging?	Yes
Give number of bootstrap runs for calculating BMD confidence interval (e.g. 200 or more)	200
Do you want to include the logit and probit model in model averaging?	No (recommended)

2.4.3. Parameter settings for benchmark dose modeling

In summary, the pre-screening steps and modeling conditions stated in Table 3 were used. The configurational setting for PROAST were set within the PROAST/RStudio software.

2.4.4. Tumorigenic Dose 50 (TD50)

TD50 was first proposed as the index of carcinogenicity potency for the original CPDB (Sawyer et al., 1984). In the assumptions that occurrence in the control group as well as intercurrent deaths are absent, the authors defined a TDx as the daily dose of chemical which gives x% of the test animals tumors by some fixed age in the control (zero-dose) group were handled as for the LD50, and intercurrent deaths were handled by life-table methods.

From the perspective of dose-response modeling, the TDx can also be modeled by the same logistic regression used previously for BMDx in eq (3).

$$TD_X = \frac{1}{b}\left(\ln\left(\frac{X}{100-X}\right) - a\right) \quad (5)$$

TD50 values given per dose-response curves in the original CPDB 2004 are based on nonparametric procedures for estimating the TD50 and for constructing confidence intervals. These are based on likelihoods which assume that the hazard is linear in dose. The original CPDB database reported calculated TD50 values for all dose-response groups along with the p-values and life table. We were not able to reproduce these original values with high confidence, presumably due to assumptions made in their methodology and also somewhat unclear descriptions in the original document. Similar observations and conclusions were also recently reported in the literature (Thresher et al., 2019).

To facilitate the use of TD50 values in further analysis, we also provide a TD50 dataset based on the CPDB 2004 database after applying the selection criteria. We applied the same inclusion criteria, i.e., used in 2012 at FDA (Aungst et al., 2012), which is roughly similar to our current study in Table 2. Excluded were all single dose studies, all studies considered "negative" or "no clear evidence", or from "all bearing animals". TD50 values were included only if findings were statistically significant (p-values $\leq$0.05).

2.4.5. Determination of a POD value

TD50 or BMDL10 are dose-response curve-specific parameters representing a defined set of effects. Per given compound, there are typically many different dose-response groups, hence many TD50 or BMD10 values available per compound, or even per given study depending on effects categories. To arrive at a value for a defined POD per given compound, the available values must be aggregated to the desired level so that the value can be reported as such.

In this study, at a given BMR of 10%, we reported minBMDL10, which is the lowest BMDL10 obtained for the given compound from all available dose-response curves that met our inclusion criteria. This minBMDL10 is then reported along with the corresponding BMD10 and BMDU10 from the same dose-response curve. The results for 421 BMDL10, along with BMD10 and BMDU10, values are provided in Supplementary Information File Table S2 (Tab "S2.2").

In the case of TD50, the same approach was followed to report the minimum of the TD50 values available from multiple dose-response curves. Supplementary Information File Table S2 (Tab "S2.4") provides minTD50 values for 616 compounds taken from the original CPDB. If the original CPDB dataset of TD50 values is considered in a POD or TTC-related analysis, minTD50 values representing specific dose-response curves for defined set of tumor descriptors should be used.

3. Results and discussion

3.1. Update of CPDB database (CPDB 2023) and new cancer TTC candidate set

The first aim of this investigation was to update rodent bioassay data published in the period of 2004–2017. The 2004 version of the legacy CPDB database (CPDB 2004) contained 1547 compounds providing 23,003 dose-response curves. The updated database (CPDB 2023) described in this paper provided a total of 1591 compounds along with 23,363 dose response curves and 6227 studies, where a study is defined by a set of common study parameters: compound, species, sex, route, dosage regimen, exposure time, citation. This effort added new data consisting of 44 compounds, 158 studies and 359 dose-response curves. Through the effort reported in this paper, rodent bioassay data for additional 59 (44 new) compounds with an additional 79 studies from EFSA, ECHA, and NTP. This new content is available in Supplementary Files Tables S1 and S2 (Tabs "S1.1" and "S2.1").

The second aim was to provide a dataset suitable for TTC analysis following an update and reanalysis of the CPDB 2004. This investigation was not intended in itself to provide updated TTC values, but to construct a new dataset to support cancer TTC. New POD datasets providing minimum BMDL10 and minimum TD50 (derived from the CPDB 2004) are provided in Supplementary File Table S2 (Tab "S2.2").

The reader is referred to Batke et al. (2021) and Escher et al. (2023) for further information on this topic. The study by Batke et al. (2021) provided a comparison of different PoDs for non-genotoxic carcinogens using an interim BMDL10 dataset, which differs from the final dataset published here, e.g. with regard to the study selection criteria and the calculation of BMD values and their confidence intervals. Establishing the new dataset suitable for TTC analysis is described below with additional details available in Supplementary Files Table S2 (Tabs "S2.1" – "S2.5").

3.2. Chemical group analysis for the new cancer TTC candidate set

The final cancer TTC candidate set consists of 421 compounds selected from the dose-response modeling evaluations described in Section 2.4. Chemicals classified as belonging to the cohort of concern (COC) were identified when applying the TTC categories available from the public ChemoTyper [https://chemotyper.org/; https://toxprint.org/]. COC chemicals in the dataset were 32 nitroso compounds, four aflatoxin-like compounds, two polychlorinated dibenzodioxins and polychlorinated dibenzofurans, and five steroids.

The result of further structural classifications is displayed in Fig. 1. To characterize the chemical structure space, compound classifications and frequency distributions were analyzed by ToxPrint chemotypes (Yang et al., 2015) against both final and interim datasets. The most abundant classes were aromatic amines, halides (aromatic/alkyl/aliphatic), and alkyl/aliphatic carboxylic acid/esters. Chemical structure spaces, as represented by ToxPrints, are very similar for both datasets. The final BMD dataset still contained structure groups assigned as COCs except the azoxy compounds (as no azoxy compounds were in the dataset).

3.3. Characterization of study data

The full updated CPDB 2023 consists of 1591 compounds providing 6227 studies in carcinogenicity, chronic toxicity, and combined chronic/carcinogenicity studies in rodents, primates, and dogs. The profile of the studies with respect to the covered species and routes of exposure is shown in Fig. 2.

The new database contains 104,911 dose-level findings (97,864 legacy and 7047 newly-added records). Considering only "positive" and "clear evidence" experimental calls, the top three most sensitive sites are liver, lung and stomach/forestomach, as indicated in Fig. 3-A. For the BMDL10 Dataset, the top sites are liver, kidney, stomach/forestomach, and followed by thyroids (Fig. 3-B).

The findings of the new database resulted in 548 types of tumors were in 248 various tissues. In Fig. 4, most prevalent tumors observed in liver, kidney and stomach/forestomach are presented for the BMDL10 Dataset. As expected from the target sites, the most prevalent and sensitive tumor types were hepatocellular carcinomas and adenomas for both the full CPDB database and the BMDL10 Dataset.

3.4. Minimum BMDL10

Following application of the inclusion criteria and assessment for BMD calculations, the final Cancer TTC dataset suitable for BMD analysis consists of 421 compounds. A POD value per given compound is determined by the minimum BMDL10 value from multiple dose response curves meeting the criteria. Supplementary Information File

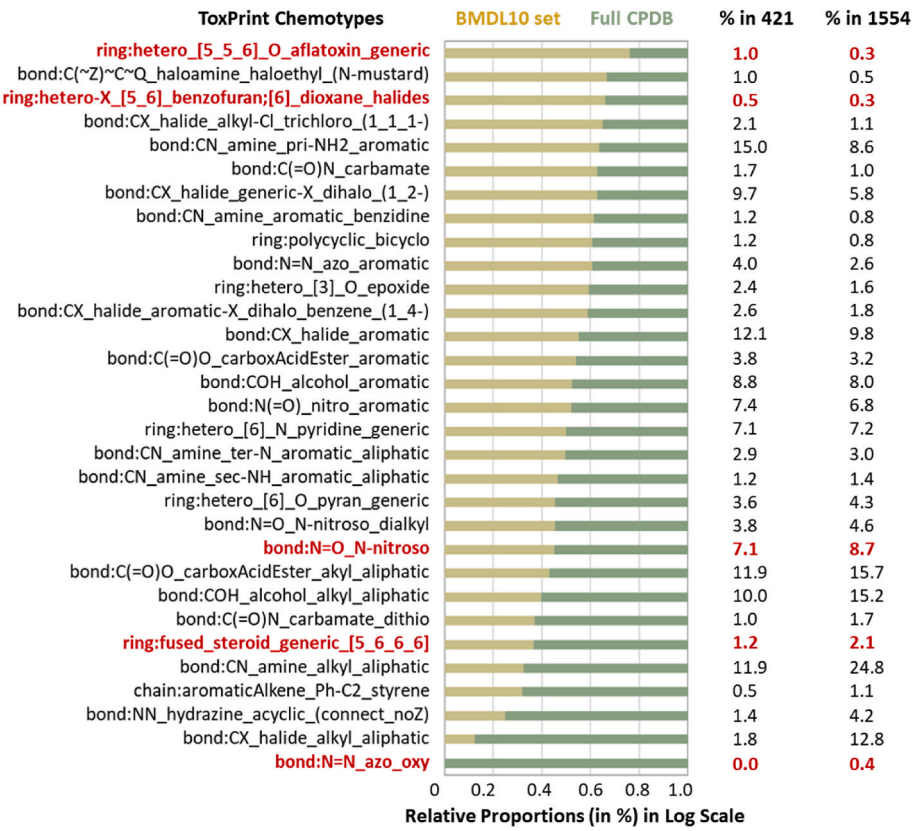

Fig. 1. Frequency of structural fragments in the full updated CPDB database and the new cancer TTC candidate set. The five Cohort of Concern categories are highlighted in red. Of the 1591 new compounds listed, 1554 compounds were structurally represented. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

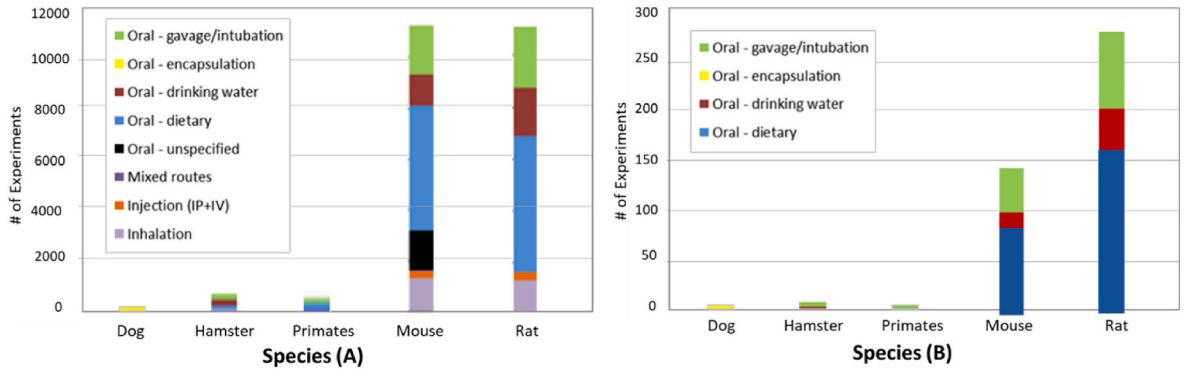

Fig. 2. Profile of the toxicity studies included in the updated CPDB (A) and the final BMDL10 Dataset (B).

Table S2 (Tabs "S2.1" – "S2.5") provide the compounds names, IDs, selected study designs, tumor descriptors, carcinogenicity calls (opinions), BMD information (minimum BMDL10, BMDU10, BMD10) and data sources.

The distribution of the minBMDL10 data is plotted in Fig. 5. This dataset will serve as the basis for further derivation of new Cancer TTC thresholds. The minBMDL10 values from both interim and final modeling results are tabulated in Supplementary Information File Table S2 (Tab "S2.2"). The interim minBMDL10 values included in Batke et al. (2021) and Escher et al. (2023) were similar to this distribution except that the quantile (natural) values of the final minBMDL10 were 2–3 times higher than the interim in the 10–95% quantile range, compared to the histogram in Fig. 5: median = 5.5; 25% quantile = 0.51; 10 % = 0.077, 5% = 0.023 mg/kg-bw/day (for N = 501). The geometric mean of the final dataset is 2-fold higher than that of the interim set, i.e., 7.9 vs. 3.6 mg/kg-bw/day.

3.5. QC analysis for BMDL10 values to establish a final dataset

For the 421 compounds in common in the interim and final sets, the minBMDL10 values varied significantly as shown in Fig. 6A. In the cases where the two datasets used the same dose-response curves, the agreement between results was much better (Fig. 6B). For a given dose-response curve, slight differences in modeling parameters generally result in relatively minor variations. The larger variations observed in Fig. 6A are thus primarily due to cases in which the reported minBMDL10 values for a given compound were based on different dose response curves in the interim and final sets. For the final dataset, we

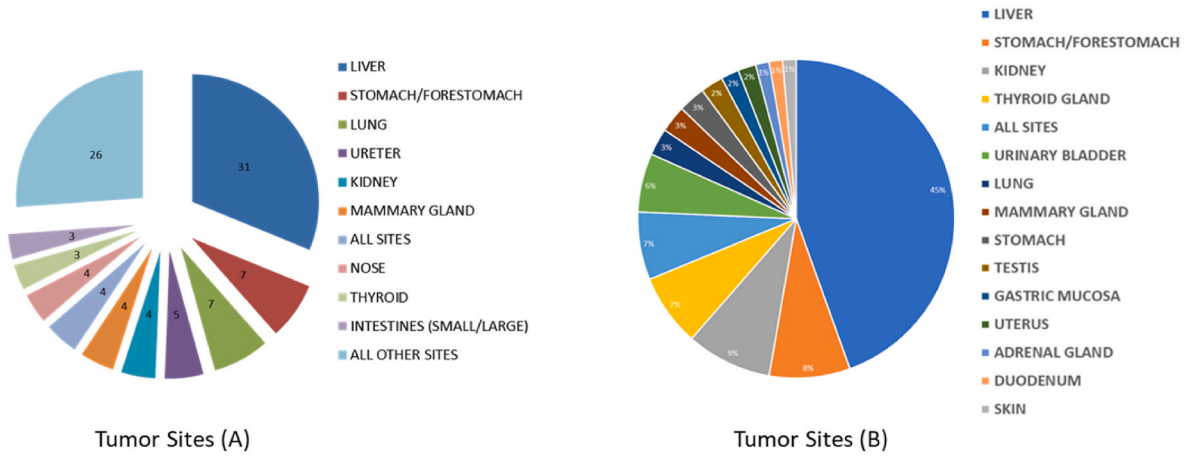

Fig. 3. The most sensitive target sites in the updated CPDB database (A) and the final BMDL10 Dataset as indicated by the percentage of records with recorded neoplastic findings in experiments with "positive" and "clear evidence" calls.

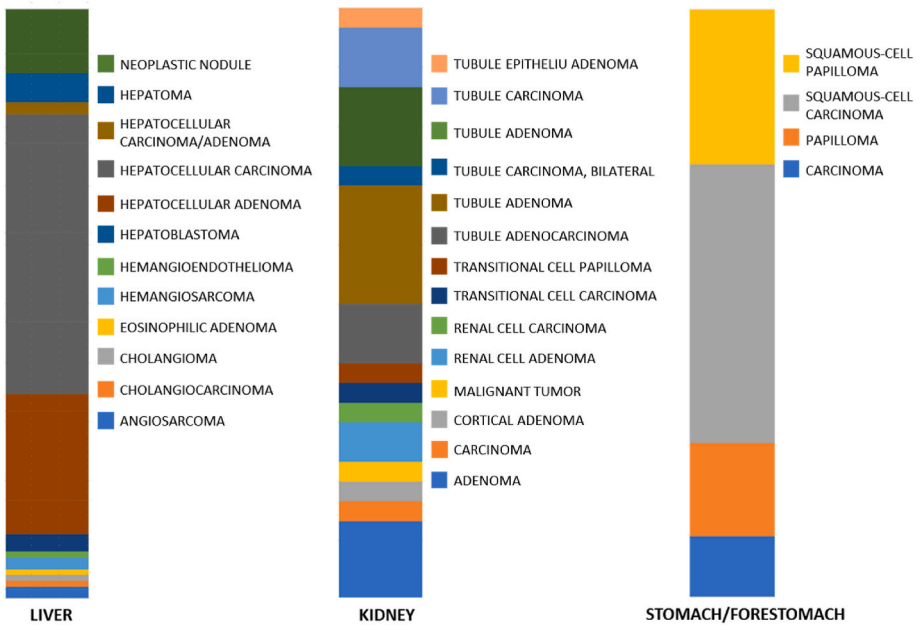

Fig. 4. Example of tumors observed at the most sensitive target sites in the liver, kidney and stomach/forestomach in the final BMDL10 Dataset in experiments with "positive" and "clear evidence" calls.

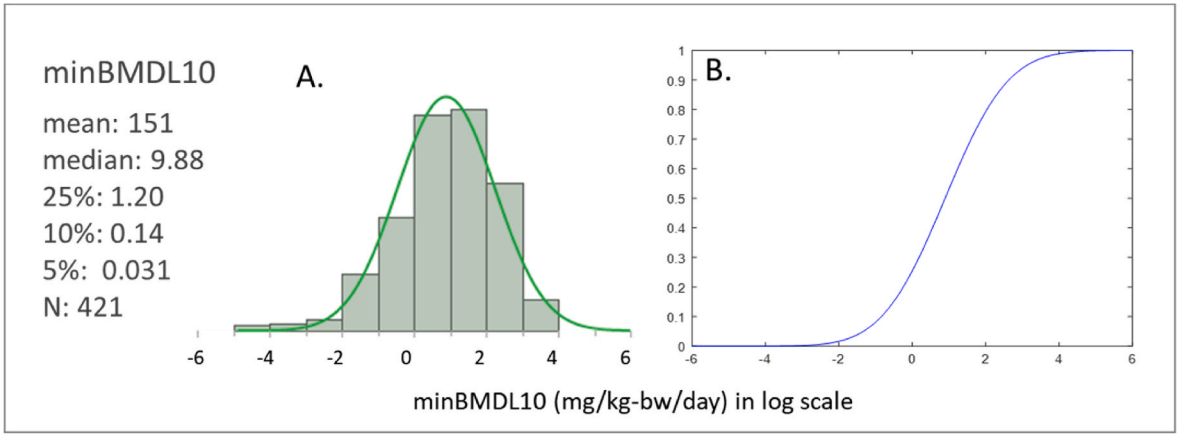

Fig. 5. Distribution of minimum BMDL10 results for the potential cancer TTC dataset. A: Histogram of minBMDL10; B: Cumulative Distribution Function of minBMDL10.

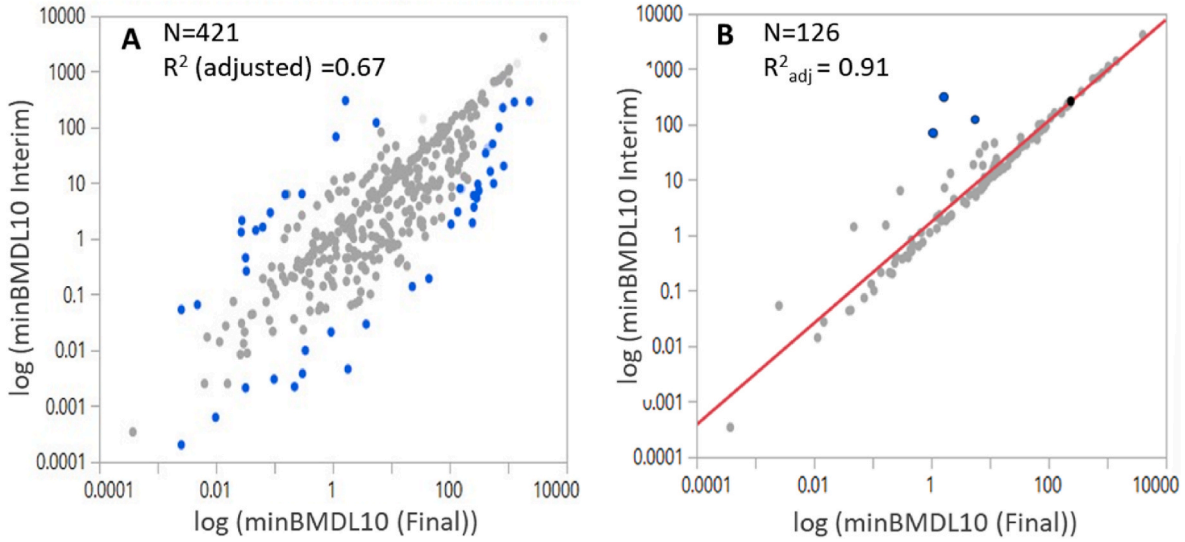

Fig. 6. Bivariate Fit of matched pairs of minBMDL10 from Interim and Final Datasets; (A) from all dose-response curves in the common set; (B) from the same dose-response curves; Blue points denote the data selected for QC. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

included additional criteria of linearity as originally recommended from CPDB publications (Sawyer et al., 1984; Thresher et al., 2019). This means that in final algorithm, a pre-screening step was added to confirm whether the logit model in eq (1) fit the data adequately before proceeding with the PROAST procedure. PROAST then evaluates multiple models and uses a model averaging method to obtain the reported BMDL10 values.

We then ran a QC step to understand the patterns of discrepancies between the interim and final sets. The dose-response curves for QC were selected systematically based on the bivariate fit analysis shown in Fig. 6 (the blue points). Interestingly, more matched pairs seem to be exhibiting larger minBMDLs from the final set; on the other hand, for the pairs based on the same dose-response curves, most deviations had higher minBMDL10 values in the interim dataset. To understand these patterns, the dose-response curves of blue-highlighted points in Fig. 6 were manually evaluated. We found that some of the interim results were derived from less-than-optimal dose-response curves including transient responses and/or high tumor findings for the control. Examples for three chemicals are compared in Fig. 7.

The results from the interim dataset had a tendency to tolerate studies with high tumor counts at control doses as well as transient dose responses. In the case of CMS-1022, using the analysis presented here for the final dataset, none of the dose-response curves satisfied the criteria for AIC (as defined in Sections 2.4.1 & 2.4.2) and BMD >0. For CMS-1447 case, both modeling runs selected the same study, except the different tumor types, i.e., liver hemangiosarcoma for the final set whereas the lung alveolar adenoma for the interim set. This particular dose response curve was rejected in the final modeling set due to the

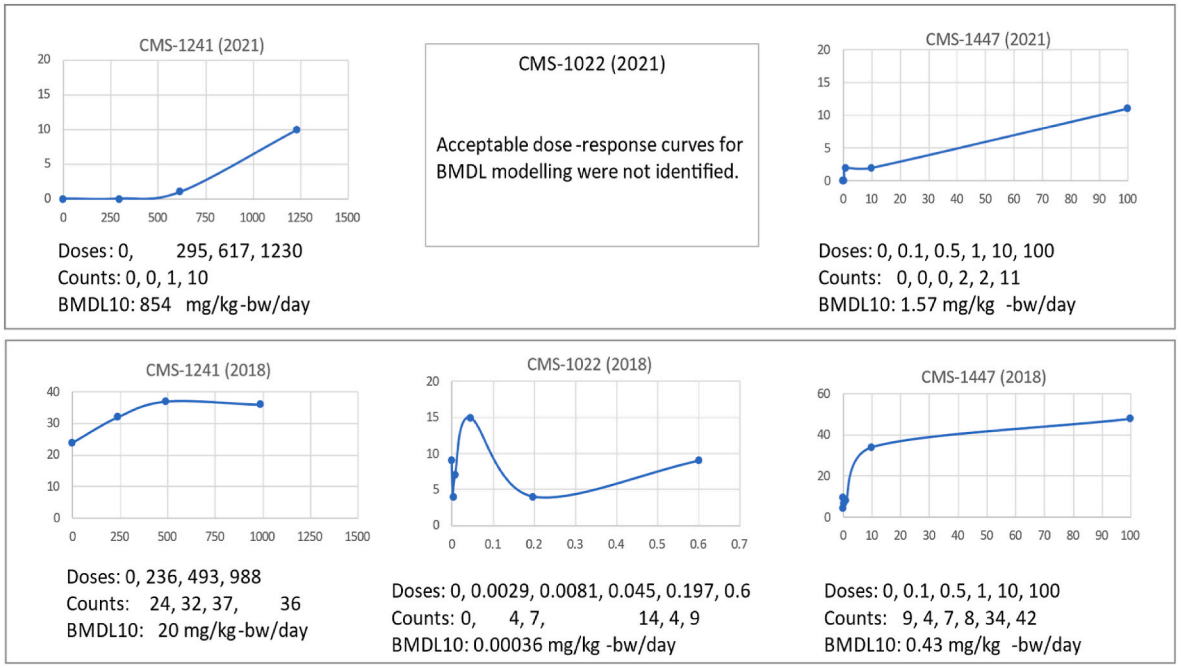

Fig. 7. Comparison of dose response curves selected for BMDL10 modeling in the interim (bottom row, 2018) and final (top row, 2021) datasets.

additional criteria that requires BMDL < BMD < BMDU.

Another source for the variations in BMDL10 modeling was found to be the parameter settings in PROAST software. For example, the specified number of bootstrap runs for calculating BMD confidence intervals in some cases made a substantial difference in the results. As an example, for CMS-383 (decabromophenyl ether) BMDL calculations based on the same dose-response curve resulted in different BMDL10 values when varying the number bootstrap runs from 10 to 1000. Although a smooth progression or convergence was not observed, BMDL10 values in this case tend to be much higher (e.g., >400 mg/kg-bw/day) for a small number of bootstrap runs but then become smaller with less variability (e.g., 1.7 mg/kg-bw/day) at higher number of bootstraps. For this reason, for the final dataset, modeling was performed using 200 bootstrap runs for all dose-response curves, which PROAST recommends as the minimum. Exploring the effect of these parameters is recommended when model averaging. Supplementary Information File Table S2 (Tab "S2.3") gives the QC results are for example of 24 compounds including 48 dose response curves.

3.6. Comparison with minTD50 and minBMDL10 for the legacy CPDB database

Although we cannot compare the minBMDL10 and minTD50 for all chemicals in the updated database, the existing records in the original CPDB database can be still compared. Any given value for a TD50 or BMDL10 is associated with a particular dose-response curve. In Fig. 8, all minimum TD50 values per given compound from all dose-response curves meeting $p \leq 0.05$ are compared with all minBMDL10 values determined in the final BMDL10 dataset.

The correlation of minBMDL10 values with the TD50 values are shown in Fig. 8A for 386 common compounds with the matching dose-response curves. In Fig. 8B, minBMDL10 values for 392 compounds were compared against the minTD50 value per given compound. It is not surprising that the two PODs are more closely correlated if values are based on the same dose-response curves as in Fig. 8A. These results also show that minTD50 values were approximately 35–350 mg/kg-bw/day greater than the minBMDL10 values at 95% confidence level.

The original CPDB provided TD50 values for each dose-response curve for all observations in a study per given compound. Similar to the selection of minimum BMD values, the minimum TD50 per compound was also determined for each compound. It is worthwhile noting that the original CPDB data are also associated with harmonic TD50 values, which is an aggregated value for multiple TD50 values to a compound or species-level across all observations involving multiple studies. A harmonic mean TD50 is therefore not tied to a particular dose-response curve, but only to a compound and/or species (e.g., rat TD50). For these reasons, harmonic mean TD50 values reported in the CPDB database or other data sources should not be used as PODs in TTC approaches. The minimum TD50 values are only reported in the TTC-related publications (Cheeseman et al., 1999; Aungst et al., 2012) and have not been publicly distributed historically.

4. Conclusions

The current study updated the legacy CPDB database and provided POD values to further enable TTC-related analysis. Addition of data based on 44 structurable compounds with new studies meeting the study inclusion criteria constitute as an update, resulting in a total compound count of 1591 and 23,633 dose-response curves. The full data of this new database were analyzed to establish a POD dataset by applying a set of rigorous POD criteria. Based on dose-response curve modeling using the publicly free PROAST/RStudio software, POD values include BMR modeling results (minBMDL10, BMD10, BMDU10) for the entire updated CPDB have been determined. Upon applying the pre-screening rules and algorithms, the modeling results yielded a dataset of 421 compounds with BMDL10 values. This final dataset is provided as a cancer TTC candidate set. Also presented in this study are the 616 minTD50 values from the original CPDB dataset meeting the equivalent study inclusion criteria as applied in the BMDL10 dataset (Supplementary Information File Table S2 (Tab "S2.4")). It is intended that this analysis and associated data set(s) will allow for a continued evaluation of the cancer potency TTC values and will support also other ongoing assessments such as the derivation of category specific threshold values.

Funding

Funding was received from the Cefic Long Range Research Initiative (LRi) through the B18 Project.

CRediT authorship contribution statement

Chihae Yang: Conceptualization, Writing – original draft, Methodology, Formal analysis, Investigation, Data curation, Visualization. **James F. Rathman:** Writing – review & editing, Methodology, Formal analysis, Investigation. **J. Vinicius Ribeiro:** Data curation. **Monika Batke:** Writing – review & editing. **Sylvia E. Escher:** Writing – review & editing. **James W. Firman:** Writing – review & editing. **Bryan Hobocienski:** Data curation. **Rupert Kellner:** Data curation. **Aleksandra Mostrag:** Writing – review & editing, Data curation. **Katarzyna R. Przybylak:** Writing – review & editing, Data curation. **Mark T.D. Cronin:** Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence

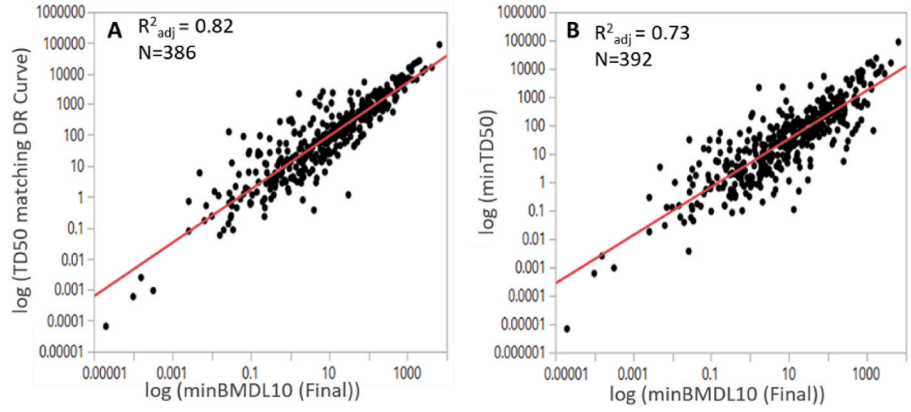

Fig. 8. Comparison of minBMDL10 and minTD50 for compounds in the original CPDB database. (DR Curve: dose-response curve).

the work reported in this paper.

Data availability

Data are available in the supplementary information files

Acknowledgements

The project received funding from the Cefic LRi LRI-B18 Project. The following are gratefully acknowledged: the knowledge and advice provided by Dr Cyril Marsaux (ILSI Europe); Drs Kirk Arvidson, Jason Aungst, and Kristi Jacobs-Muldoon (US FDA); Prof. Wout Slob (RIVM); the comments and guidance of the Organising Committee of the Cefic-LRI/ILSI Europe Joint Workshop on 'Carcinogen Dose-Response Database for Threshold of Toxicological Concern (TTC) held virtually in April 2021; Mr Sam Belfield (Liverpool John Moores University) to harvest and curate data.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.fct.2023.114182.

References

Aungst, J., Muldoon-Jacobs, K., Rua, D., Arvidson, K., Hristozov, D., Mugabe, B., Matthews, E.J., McCarthy, A., Yang, C., Cheeseman, M.A., 2012. Revisiting the TTC approach for cancer assessment: Part of a road map of computational methods at FDA CFSAN OFAS. Toxicol. Suppl. Toxicol. Sci. 126, 61.

Batke, M., Afrapoli, F.M., Kellner, R., Rathman, J.F., Yang, C., Cronin, M.T.D., Escher, S. E., 2021. Threshold of Toxicological Concern - an update for non-genotoxic carcinogens. Front. Toxicol. 3, 688321 https://doi.org/10.3389/ftox.2021.688321.

Boobis, A., Brown, P., Cronin, M.T.D., Edwards, J., Galli, C.L., Goodman, J., Jacobs, A., Kirkland, D., Luijten, M., Marsaux, C., Martin, M., Yang, C., Hollnagel, H.M., 2017. Origin of the TTC values for compounds that are genotoxic and/or carcinogenic and an approach for their re-evaluation. Crit. Rev. Toxicol. 47, 705–727. https://doi.org/10.1080/10408444.2017.1318822.

Cheeseman, M.A., Machuga, E.J., Bailey, A.B., 1999. A tiered approach to threshold of regulation. Food Chem. Toxicol. 37, 387–412. https://doi.org/10.1016/s0278-6915(99)00024-1.

CORINA Symphony, 2023. CORINA Symphony - Managing and Profiling Molecular Datasets. MN-AM, Molecular Networks GmbH, Germany-Altamira LLC, USA. https://www.mn-am.com/products/corinasymphony.

deLeeuw, J., 1992. Introduction to Akaike (1973) information theory and an extension of the maximum likelihood. In: Kotz, S., Johnson, N.L. (Eds.), Principle Breakthroughs in Statistics (Volume II): Methodology and Distribution. Springer-Verlag, New York.

Embry, M.R., Bachman, A.N., Bell, D.R., Boobis, A.R., Cohen, S.M., Dellarco, M., Dewhurst, I.C., Doerrer, N.G., Hines, R.N., Moretto, A., Pastoor, T.P., Phillips, R.D., Rowlands, J.C., Tanir, J.Y., Wolf, D.C., Doe, J.E., 2014. Risk assessment in the 21st century: roadmap and matrix. Crit. Rev. Toxicol. 44 (Suppl. 3), 6–16. https://doi.org/10.3109/10408444.2014.931924.

Escher, S.E., Felter, S.P., Hollnagel, H., Boobis, A.R., Yang, C., Rathman, J., Cronin, M.T. D., Batke, M., 2023. Workshop report on the evaluation of the updated and expanded carcinogen database to support derivation of threshold of toxicological concern values for DNA-reactive carcinogens. ALTEX - Alternat. Anim. Exp. 40, 341–349. https://doi.org/10.14573/altex.2210111.

EFSA, 2016. European Food Safety Authority (2016) Review of the Threshold of Toxicological Concern (TTC) Approach and Development of New TTC Decision Tree, vol. 13. EFSA Support. Publ.

EFSA Scientific Committee, Hardy, A., et al., 2017. Update: guidance on the use of the benchmark dose approach in risk assessment. EFSA J. 15 (4658), 41. https://doi.org/10.2903/j.efsa.2017.4658.

Gold, L.S., Manley, N.B., Slone, T.H., Garfinkel, G.B., Ames, B.M., Rohrbach, L., Stern, B. R., Chow, K., 1995. Sixth plot of the carcinogenic potency database: results of animal bioassays published in the General Literature 1989 to 1990 and by the National Toxicology Program 1990 to 1993. Environ. Health Perspect. 103, 3–122. https://doi.org/10.1289/ehp.95103s83.

Gold, L.S., Manley, N.B., Slone, T.H., Rohrbach, L., Backman Garfinkelnd, G., 2005. Supplement to the carcinogenic potency database (CPDB): results of animal bioassays published in the general literature through 1997 and by the national Toxicology Program in 1997–1998. Toxicol. Sci. 85, 747–808. https://doi.org/10.1093/toxsci/kfi161.

Gold, L.S., Sawyer, C.B., Magaw, R., Backman, G.M., de Veciana, M., Levinson, R., Hooper, N.K., Havender, W.R., Bernstein, L., Peto, R., Pike, M.C., Ames, B.N., 1984. A carcinogenic potency database of the standardized results of animal bioassays. Environ. Health Perspect. 58, 9–319. https://doi.org/10.1289/ehp.84589.

Gold, L.S., Slone, T.H., Bernstein, L., 1989. Summary of carcinogenic potency and positivity for 492 rodent carcinogens in the carcinogenic potency database. Environ. Health Perspect. 79, 259–272. https://doi.org/10.1289/ehp.8979259.

Goyak, K.O., Sarang, S.S., Franzen, A., Borghoff, S.J., Ryman-Rasmussen, J.P., 2022. Adverse outcome pathway (AOP): α2u-globulin nephropathy and kidney tumors in male rats. Crit. Rev. Toxicol. 52, 345–357. https://doi.org/10.1080/10408444.2022.2082269.

Kroes, R., Renwick, A.G., Cheeseman, M., Kleiner, J., Mangelsdorf, I., Piersma, A., Schilter, B., Schlatter, J., van Schothorst, F., Vos, J.G., Würtzen, G., 2004. Structure-based Thresholds of Toxicological Concern (TTC): guidance for application to compound present at low levels in the diet. Food Chem. Toxicol. 42, 65–83. https://doi.org/10.1016/j.fct.2003.08.006.

NIH CPDB, 2023. Download Carcinogenic Potency Database (CPDB) Data. https://www.nlm.nih.gov/databases/download/cpdb.html.

RIVM, 2023. PROAST. Available from: https://www.rivm.nl/en/proast (Last accessed April 2023).

Rulis, A.M., 1987. *De minimis* and the threshold of regulation. In: Food Protection Technology. Lewis Publishing, Chelsea, pp. 29–37.

Sawyer, C., Peto, R., Bernstein, L., Pike, M.C., 1984. Calculation of carcinogenic potency from long-term animal carcinogenesis experiments. April Biometrics 40, 27–40.

Thresher, A., Gosling, J.P., William, R., 2019. Generation of TD50 values for carcinogenicity study data. Toxicol. Res. 8, 696–703. https://doi.org/10.1039/c9tx00118b.

ToxPrint.Org. https://toxprint.org/[Last accessed in April 2023].

US FDA, 1995. Food Additives: threshold of regulation of compound used in food-contact articles: final Rule. Fed. Regist. 60, 36582–36596.

US FDA, 2000. Guidance for Industry and Other Stakeholders Toxicological Principles for the Safety Assessment of Food Ingredients (Redbook 2000), p. 286. Available from: https://www.fda.gov/media/79074/download.

Yang, C., Barlow, S.M., Muldoon Jacobs, K.L., Vitcheva, V., Boobis, A.R., Felter, S.P., Arvidson, K.B., Keller, D., Cronin, M.T.D., Enoch, S., Worth, A., Hollnagel, H.M., 2017. Thresholds of Toxicological Concern for cosmetics-related compound: new database, thresholds, and enrichment of chemical space. Fd Chem. Toxicol. 109, 170–193. https://doi.org/10.1016/j.fct.2017.08.043.

Yang, C., Cronin, M.T.D., Arvidson, K.B., Bienfait, B., Enoch, S.J., Heldreth, B., Hobocienski, B., Muldoon-Jacobs, K., Lan, Y., Madden, J.C., Magdziarz, T., Marusczyk, J., Mostrag, A., Nelms, M., Neagu, D., Przybylak, K., Rathman, J.F., Park, J., Richarz, A.-N., Richard, A.M., Ribeiro, J.V., Sacher, O., Schwab, C., Vitcheva, V., Volarath, P., Worth, A.P., 2021. COSMOS next generation – a public knowledge base leveraging chemical and biological data to support the regulatory assessment of chemicals. Comput. Toxicol. 19, 100175 https://doi.org/10.1016/j.comtox.2021.100175.

Yang, C., Tarkhov, A., Marusczyk, J., Bienfait, B., Gasteiger, J., Kleinoeder, T., Magdziarz, T., Sacher, O., Schwab, C.H., Schwoebel, J., Terfloth, L., Arvidson, K., Richard, A., Worth, A., Rathman, J., 2015. New publicly available chemical query language, CSRML, to support chemotype representations for application to data mining and modeling. J. Chem. Inf. Model. 55, 510–528. https://doi.org/10.1021/ci500667v.