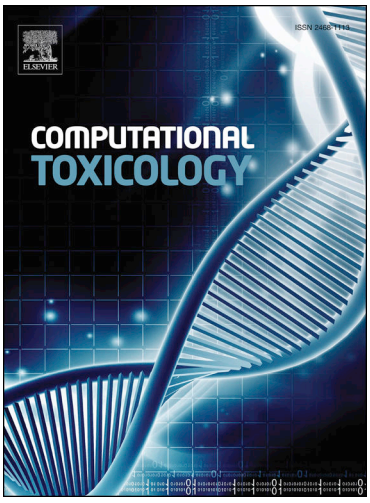
Journal Pre-proofs



The TTC Data Mart: an interactive browser for Threshold of Toxicological Concern calculations

Chantel I. Nicolas, Kevin Bronson, Salil N. Pendse, Alina Efremenko, Jeremy M. Fitzpatrick, Melyssa S. Minto, Kamel Mansouri, Miyoung Yoon, Martin B. Phillips, Rebecca A. Clewell, Melvin E. Andersen, Harvey J. Clewell III, Patrick D. McMullen

PII: S2468-1113(20)30038-4
DOI: https://doi.org/10.1016/j.comtox.2020.100128
Reference: COMTOX 100128

To appear in: *Computational Toxicology*

Received Date: 31 December 2019
Revised Date: 18 June 2020
Accepted Date: 22 June 2020

Please cite this article as: C.I. Nicolas, K. Bronson, S.N. Pendse, A. Efremenko, J.M. Fitzpatrick, M.S. Minto, K. Mansouri, M. Yoon, M.B. Phillips, R.A. Clewell, M.E. Andersen, H.J. Clewell III, P.D. McMullen, The TTC Data Mart: an interactive browser for Threshold of Toxicological Concern calculations, *Computational Toxicology* (2020), doi: https://doi.org/10.1016/j.comtox.2020.100128

The TTC Data Mart: an interactive browser for Threshold of Toxicological Concern calculations

Chantel I. Nicolas[1], Kevin Bronson[1], Salil N. Pendse[1], Alina Efremenko[1], Jeremy M. Fitzpatrick[1], Melyssa S. Minto[1], Kamel Mansouri[2], Miyoung Yoon[1], Martin B. Phillips[1], Rebecca A. Clewell[3], Melvin E. Andersen[1], Harvey J. Clewell III[4,†], Patrick D. McMullen[1,†,*]

[1]ScitoVation, LLC, Durham, NC 27713
[2]Integrated Laboratory Systems, Inc., Morrisville, NC 27560
[3]21st Century Tox Consulting, Chapel Hill, NC 27516
[4]Ramboll, 6 Davis Drive, Research Triangle Park, NC 27703
[†]These authors contributed equally to this work.

***Corresponding Author:**
100 Capitola Drive Suite 106, Durham, NC 27713
correspondingauthor@scitovation.com
Phone: (847) 477-5938

For submission to Computational Toxicology

1. Abstract

Computational technologies hold promise for informing decisions on the use of thousands of chemicals in commerce that have little or no toxicological data. Threshold of Toxicological Concern (TTC) groups compounds based on the outcome of an established analysis the chemical structure. Characterizations of hazard for representative compounds within these groups can be used in various decision contexts to evaluate risk of compounds present in trace amounts in food, drugs, and consumer products. Here, we introduce TTC calculations for a 45,000-compound chemical universe described previously in the Collaborative Estrogen Receptor Activity Prediction Project. We present these calculations in the TTC Data Mart: a web-based interface for searching and sorting compounds by TTC. The TTC Data Mart will aid in exploiting existing computational methods at the pre-assessment phase of a tiered risk-based approach to quickly, and conservatively, prioritize thousands of untested chemicals for further study.

2. Introduction

The majority of chemicals used in commerce have not undergone any formal toxicological hazard assessment. While time and resource intensive in vivo-based methods are the traditionally accepted methods for chemical safety testing, decision-makers across government and industry are now being challenged to optimize limited resources. As such, in silico and in vitro methods are being incorporated into the evaluation of the thousands of untested chemicals in commerce [1]. An attractive starting point for making risk-based decisions is the use of in silico tools to predict a chemical's potential for adverse health effects in specific exposure scenarios [2-4]. The use of computational technologies for high-throughput risk prioritization as a means of identifying chemicals for further testing has been demonstrated by regulatory programs such as the Endocrine Disruptor Screening Program [5], and was highlighted by the National Research Council as a critical part of the path towards 21st century risk assessment [6-8].

A variety of techniques in chemical safety decision-making use information about compound structure (e.g., physical chemistry, reactivity, and potential for interactions with biological components) to infer hazard. Formally, this type of structural information can be encoded as a (quantitative) structure-activity relationship ((Q)SAR) [9-12]. The association between structure and activity also underpins the toxicological concept of read across: toxicity information from data-rich compounds is mapped to data-poor compounds on the basis of structural similarity [13, 14]. An alternative concept—Threshold of Toxicological Concern (TTC)—is a system for establishing health-protective exposure guidelines based on a discrete set of chemical features. The premise behind TTC is that a threshold of acceptable exposure can be determined from chemical structures, even when directly applicable toxicity data are unavailable [15-18]. This approach has been used by industry and regulatory bodies (e.g., U.S. FDA and European Commission) for compounds that are not currently regulated or compounds that lack sufficient data for making risk assessment decisions [19-21].

The TTC approach assigns a compound to one of various “classes” of chemicals based on structural characteristics associated with toxicity [15]. These class definitions have been tied to a historical database of risk assessment information, allowing for the estimation of hazard level based on existing risk assessments of compounds with similar structural features [18]. Because TTC depends on the structure, it is potentially useful for informing gaps for data-poor compounds in the absence of chemical-specific toxicity data. Several tiered frameworks are emerging for integrating TTC and other computational techniques more generally into the chemical risk assessment landscape [22-24]. The RISK21 tiered risk assessment approach recommends using a TTC-based approach to prioritize “data-poor” chemicals of interest that lack appropriate estimates of hazard.

TTC has potential for forming putative estimates of chemical risk for large chemical datasets where it would be impractical to experimentally determine compound-specific hazard values. A comparison of hazard estimates using TTC and conventional points of departure

derived from in-life studies has supported this application of the technology [25]. In a previous application TTCs were evaluated for 8000 compounds and compared against exposure estimates as an estimate of risk [26]. One key conclusion of this study was that attaining sufficient confidence in results from a TTC-based approach would greatly enhance efficient risk-based prioritization. Here, we present TTC categories and values for a larger assembly of 45,000 compounds. These data lay the groundwork for large-scale estimation of quantitative measures of risk such as margin of exposure. These estimates have been made available in the TTC Data Mart: a searchable graphical interface, and are available for download.

3. Selection of a Large Chemical Dataset

We enlisted a structural database of 45,038 compounds that had been developed for the Collaborative Estrogen Receptor Activity Prediction Project (CERAPP) [11]. This set of chemicals has been developed as a representative universe of chemicals relevant to human exposure. The structures from the chemical dataset were provided in terms of simplified molecular-input line-entry system (SMILES) strings in a structural-data file (.sdf). Desalted QSAR ready structures were prepared using version 2.5 of KNIME workflow available here: https://github.com/kmansouri/QSAR-ready/releases [27]. We also used original SMILES strings taken from the CERAPP data.

The CERAPP dataset underwent a rigorous curation process wherein inorganic structures, salts, metals, and mixtures were removed to arrive at structures that are appropriate for computational chemistry approaches, such as QSARs. The QSAR-ready structures from the chemical dataset were provided in terms of simplified molecular-input line-entry system (SMILES) strings in a structural-data file (.sdf).

4. Estimation of Thresholds of Toxicological Concern

Toxtree (v3.1.0) was employed using imported QSAR-ready SMILES strings via batch processes applied with various structure-based decision trees [28] (Figure 1). The Kroes TTC Decision Tree [16] was used to identify chemicals for which TTC is not applicable. Cramer Rules with Extensions [15, 28-30] were used to assign compounds to the three original Cramer structural classes: these three structural classes are presumed to represent chemicals of Low (Class I), Intermediate (Class II), and High (Class III) toxicity. Potential genotoxic carcinogens were identified using a Benigni/Bossa Rule base for Mutagenicity and Carcinogenicity method to flag compounds that might be genotoxic based on whether the result for "Structural Alert for genotoxic carcinogenicity" was "YES", and, second, an *in vitro* mutagenicity (Ames test) alerts by ISS was applied to identify compounds that have the potential for in vitro genotoxicity based on whether the result for "Structural Alert for *S. typhimurium* mutagenicity" was "YES" [31-34]. If the chemical had a "YES" for either question was flagged as potentially genotoxic [35][36]. Further, potential for *Anti-cholinesterase* activity was determined using the structural alerts for functional group identification (ISSFUNC) method to identify compounds that were (a) "carbamate esters" or (b) "organophosphate esters" based on a "YES" for "FG52_2" or "FG81_2", respectively, for each chemical in the Toxtree output file. Structures that were found to have either functional group were labeled as cholinesterase inhibitors [29, 36, 37]. The remaining chemicals were assigned a Cramer Class with extension result. Structures producing errors in Toxtree were not assigned a class.

Each chemical classification has an associated human exposure threshold value (TTC) corresponding to the daily exposure level below which there is no increased probability of risk after a lifetime of exposure. TTCs are derived by determining the 5th percentile in the distribution of no observed effect levels (NOELs) for compounds within the class and dividing these values by a safety factor of 100. Cramer class I, II, and III compounds were assigned TTCs of 30, 9.1,

and 1.5 µg/kg/d, respectively [18]. Anticholinesterases are assigned a TTC value of 0.3 µg/k/d, which was determined from in vivo data for compounds with the same mode of action and is more conservative than the Cramer class III [29, 37]. Compounds flagged as genotoxic, are assigned a TTC value of 0.0025 µg/kg/day as described previously [38, 39]. This is the lowest TTC value and is designed to be a highly conservative estimate of potential cancer risk, corresponding to an estimated excess risk of 1 in 10^6 for lifetime cancer. This qualitative difference in how TTC values are defined complicates interpretations based on direct comparisons between genotoxic and non-genotoxic classes.

Of the 45,038 compounds considered here, 66.1% received a Cramer classification, 0.6% had a structural alert for anticholinesterase activity, and 21.9% had structural alerts for genotoxicity (Figure 2). The balance of compounds were not assigned a class (Figure 1, “TTC Not Applicable”). This breakdown is consistent with a previous observation on a large chemical cohort, which found 62.8% with a Cramer classification, 1.3% with anticholinesterase alerts, and 24.1% with genotoxicity alerts [40]. Interestingly, another analysis focusing on compounds with IRIS risk assessments also reflected some similarities in the distribution [41], although with slightly higher frequency of assignments to Cramer Class III and the AChE and Genotoxicity Alert categories.

5. Development of the TTC Data Mart

TTC calculations are stored in an SQLite database. The Data Mart interface is coded in R using the Shiny package (https://shiny.rstudio.com). The TTC Data Mart can be accessed at https://scitovation.shinyapps.io/TTCApplet/ (**Error! Reference source not found.**). The TTC Data Mart presents a point-and-click table interface wherein chemical entities are displayed as rows and their properties (name, TTC, etc.) are displayed as columns. Compounds can be filtered based on TTC class and other flags, such as the existence of suitable data for performing high-throughput in vitro-to-in vivo extrapolation [42, 43] or the existence of reference

dose/concentration values in US EPA's Integrated Risk Information System. Data corresponding to a query can be downloaded as a text file suitable for import into a spreadsheet application. The database and source for the TTC Data Mart application are also available for direct download at https://github.com/ScitoVation/ttcdatamart.

6. Example: Determining Margin of Exposure for EPA Priority Candidates

The TTC Data Mart provides accessible TTC values for comparison to exposure estimates [40], in vitro points of departure [25], and established regulatory levels [41]. As an example of such a comparison, we considered the relationship between TTC values and exposure estimates for a set of compounds with presumably low levels of toxicity. The Toxic Substances Control Act requires the US Environmental Protection Agency to classify chemicals currently used in commerce as high- or low-priority according to the need to execute a formal risk evaluation. Twenty chemicals selected from the EPA's Safer Chemicals Ingredients List were designated "Low-Priority", indicating that they are generally understood to be of low risk. Chemical risk is often expressed as a margin of exposure: the ratio of a compound's estimated hazard versus its anticipated exposure. A margin of exposure less than one (i.e., a compound for which anticipated exposure exceeds its hazard level) would indicate high priority. In contrast, one expects the low priority TSCA chemicals to have large margins of exposure. Exposure estimates were available for 16 of the 20 low priority TSCA chemicals in the EPA CompTox Chemistry Dashboard [44]. Here, we define the margin of exposure for a compound as the ratio of the compound's TTC value and its exposure estimate (Table 1),

$$MoE = \frac{TTC}{\substack{Estimated \\ Exposure}} .$$

MoE was calculated using two estimates of exposure: the estimated population median exposure and the estimated 95th-percentile of the population exposures. Characteristic (geometric mean)

margin of exposure is 5.0×10^3 for the population median exposure and 55 for the 95th-percentile of exposure estimates. The smallest margin of exposure computed was 3.94, for 1-octadecanol. It should be noted that though this margin of exposure is near unity, it includes the 100-fold uncertainty factor propagated from the TTC value (i.e., the TTC is 100-fold lower than the 5^{th}-percentile of the distribution of NOAEL values). This exercise reveals that these compounds, as expected, have a large margin of exposure, and demonstrates the utility for TTC to provide a conservative proxy for hazard for chemicals that lack traditional in vivo toxicity endpoints.

7. Acknowledgements

This work was supported by the American Chemistry Council's Long-range Research Initiative. The authors thank Richard A. Becker for helpful discussions.

Table 1. Cramer classification, TTC, and inferred margin of exposure for 16 substances listed as low-priority by the US EPA under the Toxic Substances Control Act.

Chemical Name	CAS	Median EPA Dashboard Exposure (95%-ile, mg/kg/d)	Cramer Classification (TTC, mg/kg/d)	Margin of Exposure (TTC / Median Population Exposure)	Margin of Exposure (TTC / 95% Population Exposure)
3-Methoxybutyl acetate	4435-53-4	NA	Class III (1.5e-3)	Not Calculated	Not Calculated
Monosodium D-glucoheptonate	31138-65-5	9.06E-08 (8.55E-06)	TTC Not Applicable	Not Calculated	Not Calculated
D-Gluconic acid	526-95-4	6.55E-07 (8.03E-05)	Class III (1.5e-3)	2.29E+03	1.87E+01
Calcium D-gluconate	299-28-5	1.26E-06 (9.32E-05)	Class III† (1.5e-3)	1.19E+03	1.61E+01
Gluconolactone	90-80-2	1.60E-07 (1.88E-05)	Class III† (1.5e-3)	9.38E+03	7.98E+01
Potassium D-gluconate	527-07-1	1.26E-06 (9.32E-05)	Class III† (1.5e-3)	1.19E+03	1.61E+01
Sodium D-gluconate	527-07-1	6.80E-07 (7.39E-05)	Class III† (1.5e-3)	2.21E+03	2.03E+01
Dibutyl decanedioate	109-43-3	9.35E-08 (1.04E-05)	Class I (3.0e-02)	3.21E+05	2.88E+03
1-Docosanol	661-19-8	8.44E-07 (8.75E-05)	Class I (3.0e-02)	3.55E+04	3.43E+02
1-Eicosanol	629-96-9	NA	Class I (3.0e-02)	Not Calculated	Not Calculated
1-Octadecanol	112-92-5	9.53E-05 (7.62E-03)	Class I (3.0e-02)	3.15E+02	3.94E+00
Tripropylene glycol butyl ether	55934-93-5	7.00E-07 (9.57E-05)	Class III (1.5e-3)	2.14E+03	1.57E+01
Diethyl propanedioate	105-53-3	3.64E-06 (2.13E-04)	Class I (3.0e-02)	8.24E+03	1.41E+02
Dimethyl malonate	108-59-8	3.72E-07 (3.37E-05)	Class I (3.0e-02)	8.06E+04	8.90E+02
1,1'-Oxybis-2-propanol	110-98-5	3.95E-06 (2.67E-04)	Class III (1.5e-3)	3.80E+02	5.62E+00
Squalane	111-01-3	3.53E-06 (2.34E-04)	Class I (3.0e-02)	8.50E+03	1.28E+02

* Exposure estimates are based on potassium salt
† based on the TTC of the corresponding acid

8. References

1. USEPA, *Assessing and Managing Chemicals under TSCA: The Frank R. Lautenberg Chemical Safety for the 21st Century Act.* 2017.
2. Moreau, M., et al., *Using exposure prediction tools to link exposure and dosimetry for risk-based decisions: A case study with phthalates.* Chemosphere, 2017. **184**: p. 1194-1201.
3. Wambaugh, J.F., et al., *High throughput heuristics for prioritizing human exposure to environmental chemicals.* Environ Sci Technol, 2014. **48**(21): p. 12760-7.
4. Wambaugh, J.F., et al., *High-throughput models for exposure-based chemical prioritization in the ExpoCast project.* Environ Sci Technol, 2013. **47**(15): p. 8479-88.
5. USEPA, *Use of high throughput assays and computational tools; Endocrine Disruptor Screening Program; Notice of availability and opportunity for comment (EPA-HQ-OPPT-2015-0305; FRL–9928–69).* 2015.
6. NRC, *Toxicity Testing in the 21st Century: A Vision and a Strategy.* 2007, Washinton D.C.: National Academy Press.
7. NRC, *Exposure Science in the 21st Century: A Vision and A Strategy.* 2012, Washington D.C.: National Academy Press.
8. NRC, *Using 21st Century Science to Improve Risk-Related Evaluations.* 2017, Washington D.C.: National Academy Press.
9. Hansch, C., et al., *Quantitative structure-activity relationships of antineoplastic drugs: nitrosoureas and triazenoimidazoles.* Cancer Chemother Rep, 1972. **56**(4): p. 443-56.
10. Hansch, C., et al., *The expanding role of quantitative structure-activity relationships (QSAR) in toxicology.* Toxicol Lett, 1995. **79**(1-3): p. 45-53.
11. Mansouri, K., et al., *CERAPP: Collaborative Estrogen Receptor Activity Prediction Project.* Environ Health Perspect, 2016. **124**(7): p. 1023-33.
12. Cherkasov, A., et al., *QSAR modeling: where have you been? Where are you going to?* J Med Chem, 2014. **57**(12): p. 4977-5010.
13. European Chemicals Agency, *Read-Across Assessment Framework (RAAF)*. 2017.
14. Patlewicz, G., et al., *Navigating through the minefield of read-across tools: A review of in silico tools for grouping.* Computational Toxicology, 2017. **3**: p. 1-18.
15. Cramer, G.M., R.A. Ford, and R.L. Hall, *Estimation of toxic hazard--a decision tree approach.* Food Cosmet Toxicol, 1978. **16**(3): p. 255-76.
16. Kroes, R. and G. Kozianowski, *Threshold of toxicological concern (TTC) in food safety assessment.* Toxicol Lett, 2002. **127**(1-3): p. 43-6.
17. Hartung, T., *Thresholds of Toxicological Concern: Setting a threshold for testing below which there is little concern.* Alternatives to animal experimentation: ALTEX, 2017. **34**(3): p. 331-351.
18. Munro, I.C., et al., *Correlation of structural class with no-observed-effect levels: a proposal for establishing a threshold of concern.* Food Chem Toxicol, 1996. **34**(9): p. 829-67.
19. Feigenbaum, A., et al., *Reliability of the TTC approach: learning from inclusion of pesticide active substances in the supporting database.* Food Chem Toxicol, 2015. **75**: p. 24-38.
20. Leeman, W.R., L. Krul, and G.F. Houben, *Reevaluation of the Munro dataset to derive more specific TTC thresholds.* Regul Toxicol Pharmacol, 2014. **69**(2): p. 273-8.
21. Yang, C., et al., *Thresholds of Toxicological Concern for cosmetics-related substances: New database, thresholds, and enrichment of chemical space.* Food Chem Toxicol, 2017. **109**(Pt 1): p. 170-193.

22. Thomas, R.S., et al., *The next generation blueprint of computational toxicology at the U.S. Environmental Protection Agency.* Toxicol Sci, 2019.
23. Andersen, M., *Developing context appropriate toxicity testing approaches using new alternative methods (NAMs).* Altex, 2019: p. 532-534.
24. Embry, M.R., et al., *Risk assessment in the 21st century: roadmap and matrix.* Critical reviews in toxicology, 2014. **44**(sup3): p. 6-16.
25. Paul Friedman, K., et al., *Utility of In Vitro Bioactivity as a Lower Bound Estimate of In Vivo Adverse Effect Levels and in Risk-Based Prioritization.* Toxicological Sciences, 2019.
26. Patlewicz G, W.J., Felter S, Simon TW, Becker RA. , *Utilizing Threshold of Toxicological Concern (TTC) with High Throughput Exposure Predictions (HTE) as a Risk-Based Prioritization Approach for thousands of chemicals.* Computational Toxicology 2018. **in press**.
27. Mansouri, K., et al., *An automated curation procedure for addressing chemical errors and inconsistencies in public datasets used in QSAR modelling.* SAR and QSAR in Environmental Research, 2016. **27**(11): p. 911-937.
28. Patlewicz, G., et al., *An evaluation of the implementation of the Cramer classification scheme in the Toxtree software.* SAR QSAR Environ Res, 2008. **19**(5-6): p. 495-524.
29. Munro, I.C., A.G. Renwick, and B. Danielewska-Nikiel, *The Threshold of Toxicological Concern (TTC) in risk assessment.* Toxicol Lett, 2008. **180**(2): p. 151-6.
30. Lapenna, S. and A. Worth, *Analysis of the Cramer classification scheme for oral systemic toxicity-implications for its implementation in Toxtree.* JRC Scientific and Technical Report EUR, 2011. **24898**.
31. Kirkland, D., et al., *Updated recommended lists of genotoxic and non-genotoxic chemicals for assessment of the performance of new or improved genotoxicity tests.* Mutation Research/Genetic Toxicology and Environmental Mutagenesis, 2016. **795**: p. 7-30.
32. Benigni, R., *The Benigni/Bossa rulebase for mutagenicity and carcinogenicity–a module of Toxtree.* JRC Scientific and Technical Reports, 2008: p. 1-78.
33. Benigni, R. and C. Bossa, *Structure alerts for carcinogenicity, and the Salmonella assay system: a novel insight through the chemical relational databases technology.* Mutation Research/Reviews in Mutation Research, 2008. **659**(3): p. 248-261.
34. Benigni, R., C. Bossa, and O. Tcheremenskaia, *Nongenotoxic carcinogenicity of chemicals: mechanisms of action and early recognition through a new set of structural alerts.* Chemical reviews, 2013. **113**(5): p. 2940-2957.
35. Cheeseman, M.A., E.J. Machuga, and A.B. Bailey, *A tiered approach to threshold of regulation.* Food Chem Toxicol, 1999. **37**(4): p. 387-412.
36. Kroes, R., et al., *Threshold of toxicological concern for chemical substances present in the diet: a practical tool for assessing the need for toxicity testing.* Food Chem Toxicol, 2000. **38**(2-3): p. 255-312.
37. Benigni, R., O. Tcheremenskaia, and A. Worth, *Computational Characterisation of Chemicals and Datasets in Terms of Organic Functional Groups-a New Toxtree Rulebase.* JRC Scientific and Technical Reports, 2011.
38. Muller, L., et al., *A rationale for determining, testing, and controlling specific impurities in pharmaceuticals that possess potential for genotoxicity.* Regul Toxicol Pharmacol, 2006. **44**(3): p. 198-211.
39. Kroes, R., et al., *Structure-based thresholds of toxicological concern (TTC): guidance for application to substances present at low levels in the diet.* Food and chemical toxicology, 2004. **42**(1): p. 65-83.

40. Patlewicz, G., et al., *Utilizing Threshold of Toxicological Concern (TTC) with High Throughput Exposure Predictions (HTE) as a Risk-Based Prioritization Approach for thousands of chemicals.* Computational Toxicology 2018. **in press**.
41. Pham, L.L., S.J. Borghoff, and C.M. Thompson, *Comparison of threshold of toxicological concern (TTC) values to oral reference dose (RfD) values.* Regul Toxicol Pharmacol, 2020. **113**: p. 104651.
42. Yoon, M., et al., *Quantitative in vitro to in vivo extrapolation of cell-based toxicity assay results.* Crit Rev Toxicol, 2012. **42**(8): p. 633-52.
43. Wetmore, B.A., et al., *Integration of dosimetry, exposure, and high-throughput screening data in chemical toxicity assessment.* Toxicol Sci, 2012. **125**(1): p. 157-74.
44. Williams, A.J., et al., *The CompTox Chemistry Dashboard: a community data resource for environmental chemistry.* Journal of Cheminformatics, 2017. **9**(1).

9. Figures

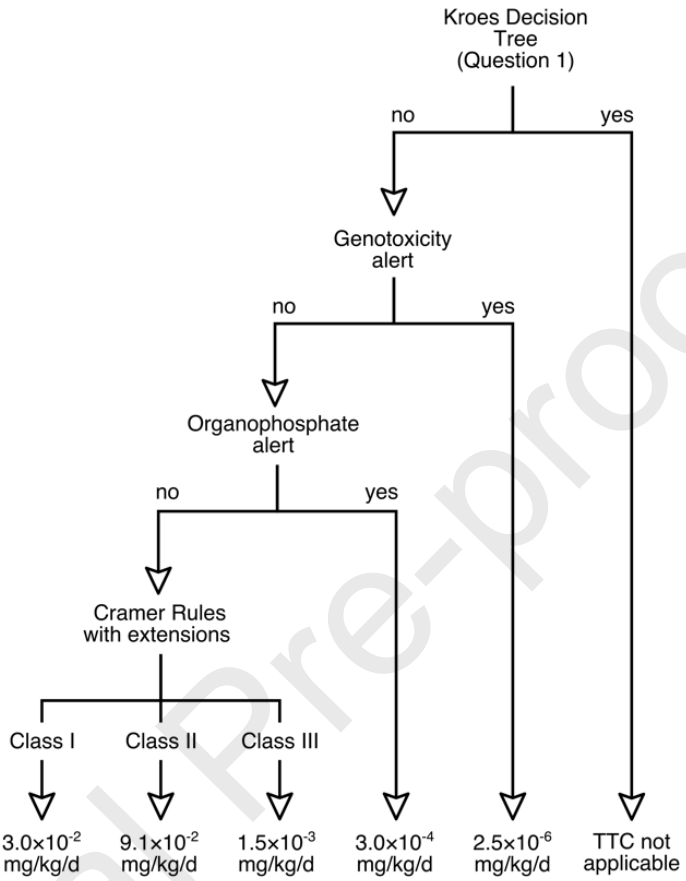


Figure 1. Workflow describing evaluation of TTCs of based on Kroes and Cramer decision trees.

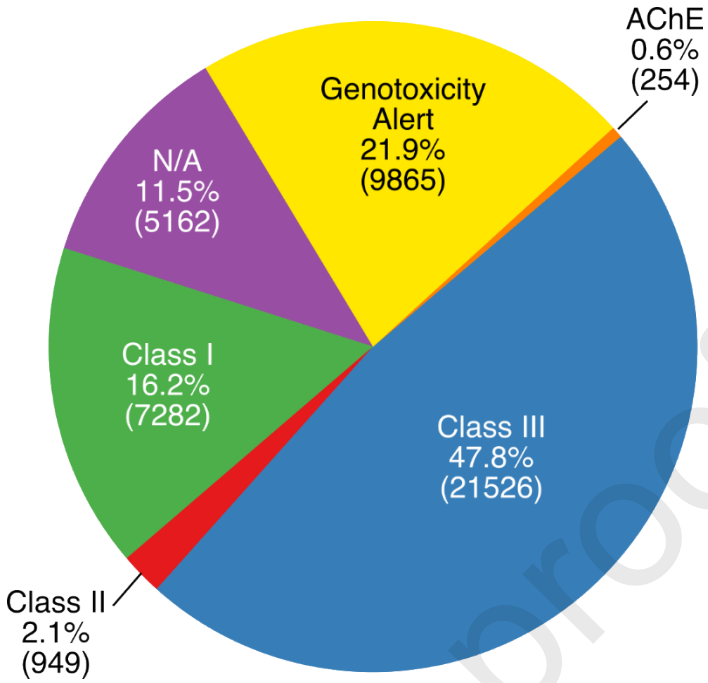


Figure 2. Distribution of compounds among the structural classes as assigned by the workflow in Figure 1.

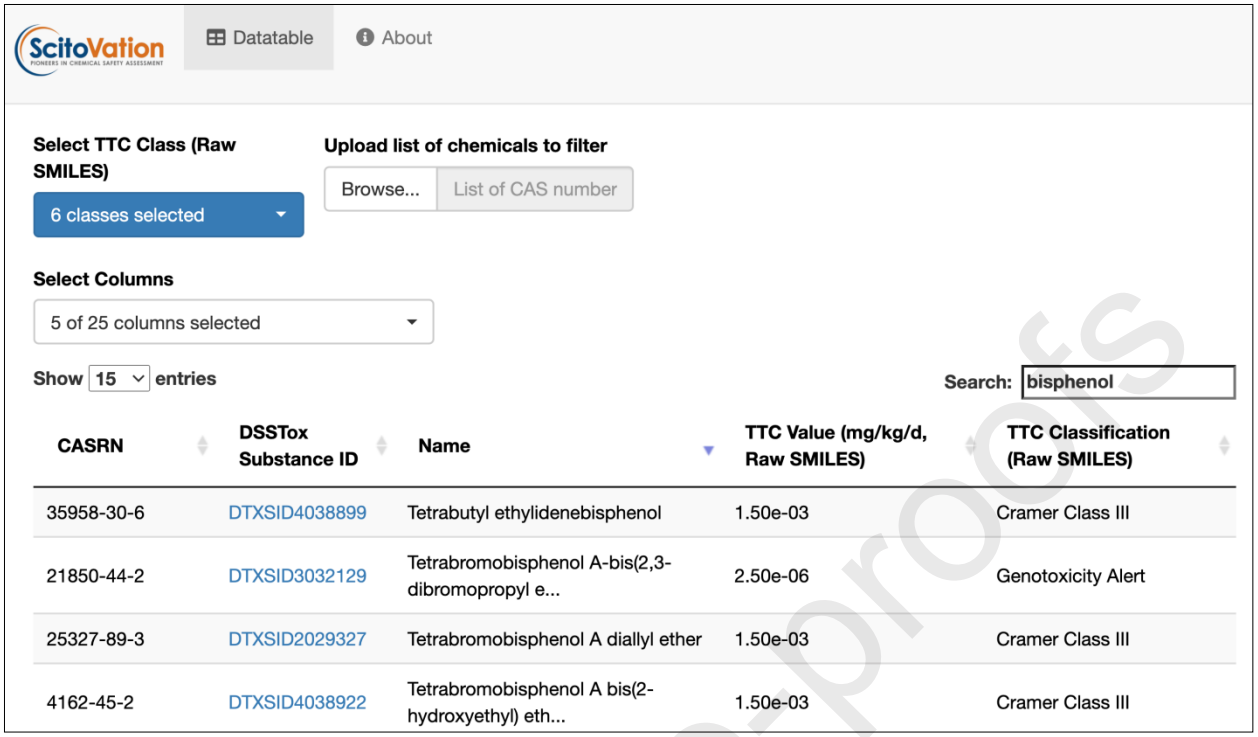


Figure 3. The TTC Data Mart interface (https://scitovation.shinyapps.io/TTCApplet/).

Chantel I. Nicolas: Methodology, Formal analysis, Investigation, Writing - Original Draft. **Kevin Bronson:** Software, Methodology, Formal analysis, Investigation. **Salil N. Pendse:** Software, Methodology, Formal analysis, Investigation. **Alina Efremenko:** Methodology, Formal analysis, Investigation. **Jeremy M. Fitzpatrick:** Methodology, Formal analysis, Investigation. **Melyssa S. Minto:** Methodology, Formal analysis, Investigation. **Kamel Mansouri:** Methodology, Formal analysis, Investigation. **Miyoung Yoon:** Conceptualization, Funding acquisition, Writing - Review & Editing.

Martin B. Phillips: Methodology, Formal analysis, Investigation. **Rebecca A. Clewell:** Conceptualization, Funding acquisition, Writing - Review & Editing. **Melvin E. Andersen:** Conceptualization, Funding acquisition, Writing - Review & Editing. **Harvey J. Clewell III:** Conceptualization, Funding acquisition, Writing - Review & Editing. **Patrick D. McMullen:** Conceptualization, Funding acquisition, Software, Writing - Review & Editing.

Research highlights

- We calculated Threshold of Toxicological Concern values for a 45,000-compound chemical universe
- We introduce an online resource called the TTC Data Mart to make these calculations accessible.
- We demonstrate how these tools can be combined to estimate margin of exposure for data-poor compounds