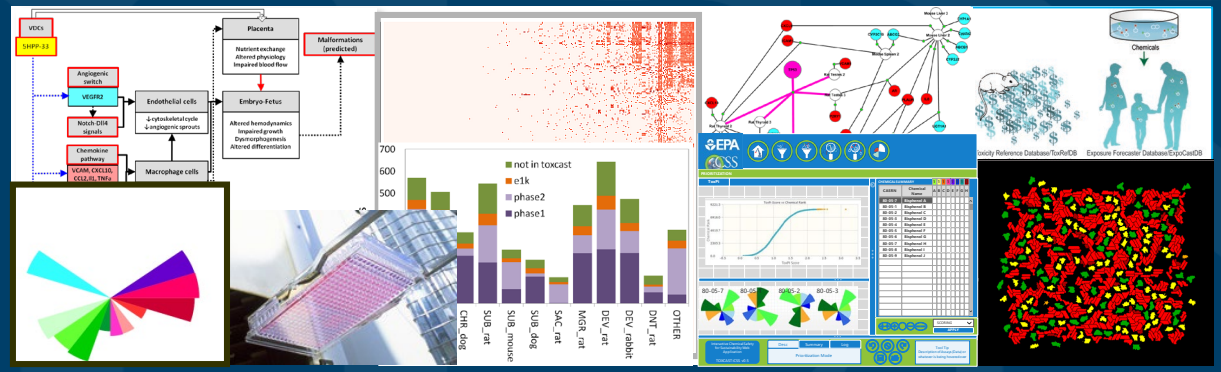
Perspectives on the Development, Evaluation, and Application of *in Silico* Approaches for Predicting Toxicity



Grace Patlewicz
Center for Computational Toxicology and Exposure (CCTE), US EPA

Conflict of Interest Statement

No conflict of interest declared.

Disclaimer:

The views expressed in this presentation are those of the authors and do not necessarily reflect the views or policies of the U.S. EPA

EPA United States Environmental Protection Agency

Outline

- Regulatory Drivers
- Computational (*in silico*) Toxicology [scope for today's talk]
- Integrated Approaches to Testing and Assessment (IATA) – definitions and Adverse Outcome Pathway (AOP) informed
- Decision contexts and their impact on the approaches applied
- Risk-based prioritisation
  - Thresholds for Toxicological Concern (TTC)
- Read-across approaches
  - Generalised Read-across (GenRA)
  - Perfluorinated & polyfluorinated substances (PFAS)
- Summary remarks
- Acknowledgements

Regulatory and Non-Regulatory drivers

- Societal demands for safer and sustainable chemical products are stimulating changes in toxicity testing and assessment frameworks
- Chemical safety assessments are expected to be conducted faster and with fewer animals, yet the number of chemicals that require assessment is also rising with the number of different regulatory programmes worldwide.
- In the EU, the use of alternatives to animal testing is promoted.
- Animal testing is prohibited in some sectors e.g. EU Cosmetics regulation
- The European Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH) legislation lays out specific information requirements, based on tonnage level triggers. However, the regulation explicitly expresses the need to use non-testing approaches to reduce the extent of experimental testing in animals.

Regulatory and Non-Regulatory drivers

- REACH-like schemes also have been established in China, South Korea, and Turkey.
- In the US, the new Frank Lautenberg Chemical Safety for the 21st Century Act (LCSA) requires that a risk based prioritisation is conducted for all substances in commerce, ~40,000, many of which are lacking sufficient publicly available toxicity information.
- EPA Administrator signed memo 10/9/19 to "direct the agency to aggressively reduce animal testing, including reducing mammal study requests and funding 30% by 2025 and completely eliminating them by 2035"
- Risk based prioritisation is also an important aspect of regulatory frameworks in Canada (the Domestics Substance List), Australia and the EU.
- Non-testing approaches offer a means of facilitating the regulatory challenges in chemical safety assessment

Computational (*In Silico*) Toxicology

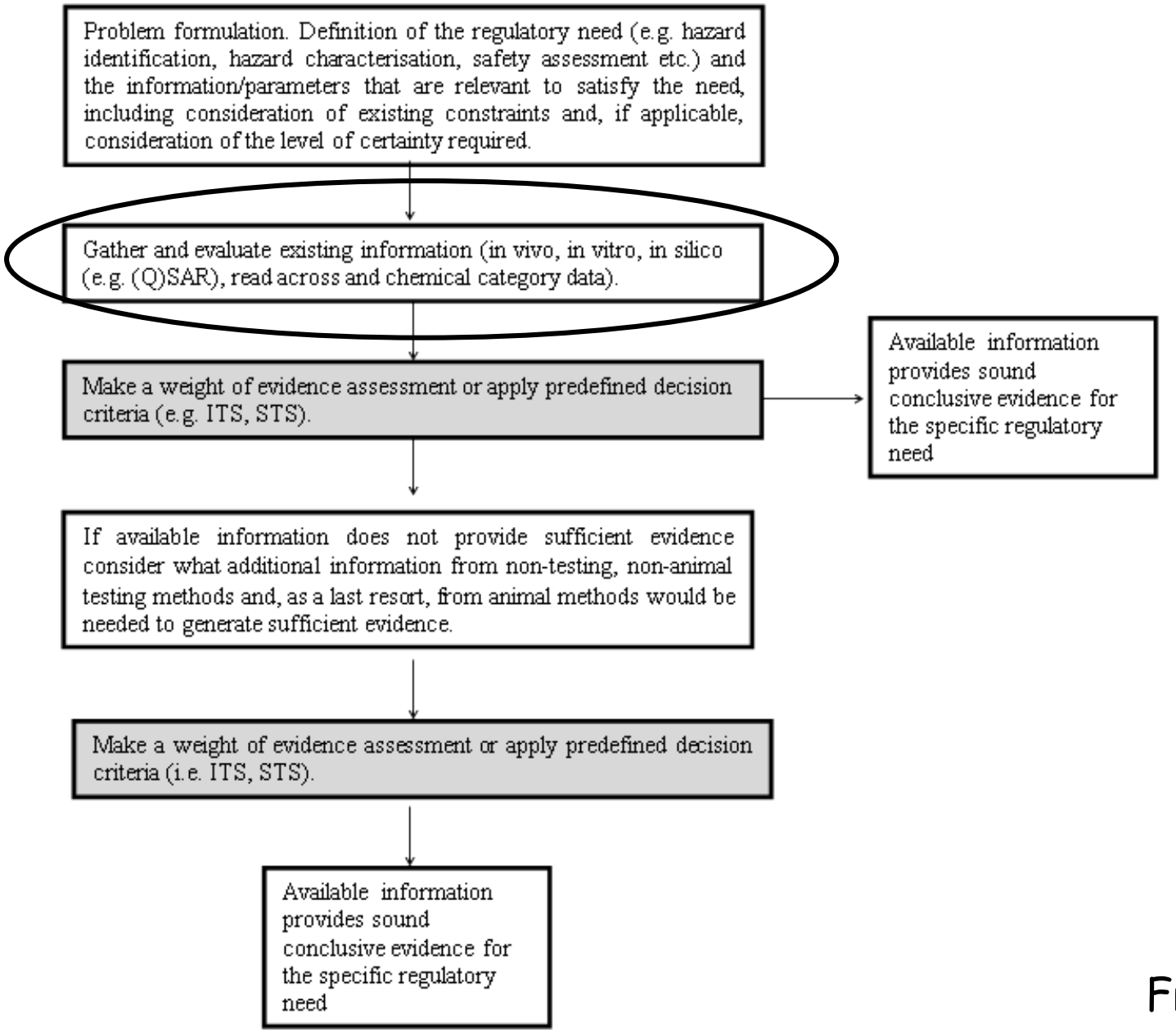
- Databases/Dashboards of existing inf
- Structure-
- Quantitativ
- Expert Sys
- Category fo
- Bioinformatics
- Chemoinformatics
- Biokinetics (PBPK)

Integrated Approaches to Testing and Assessment (IATA)

- "Integrated Testing Strategies (ITS) are …. approaches that integrate different types of data and information into the decision-making process. …"

- "A tiered approach to data gathering, testing, and assessment that integrates different types of data (including physicochemical and other chemical properties as well as *in vitro* and *in vivo* toxicity data). When combined with estimates of exposure in an appropriate manner, the IATA provides predictions of risk."

General framework of an IATA



From OECD

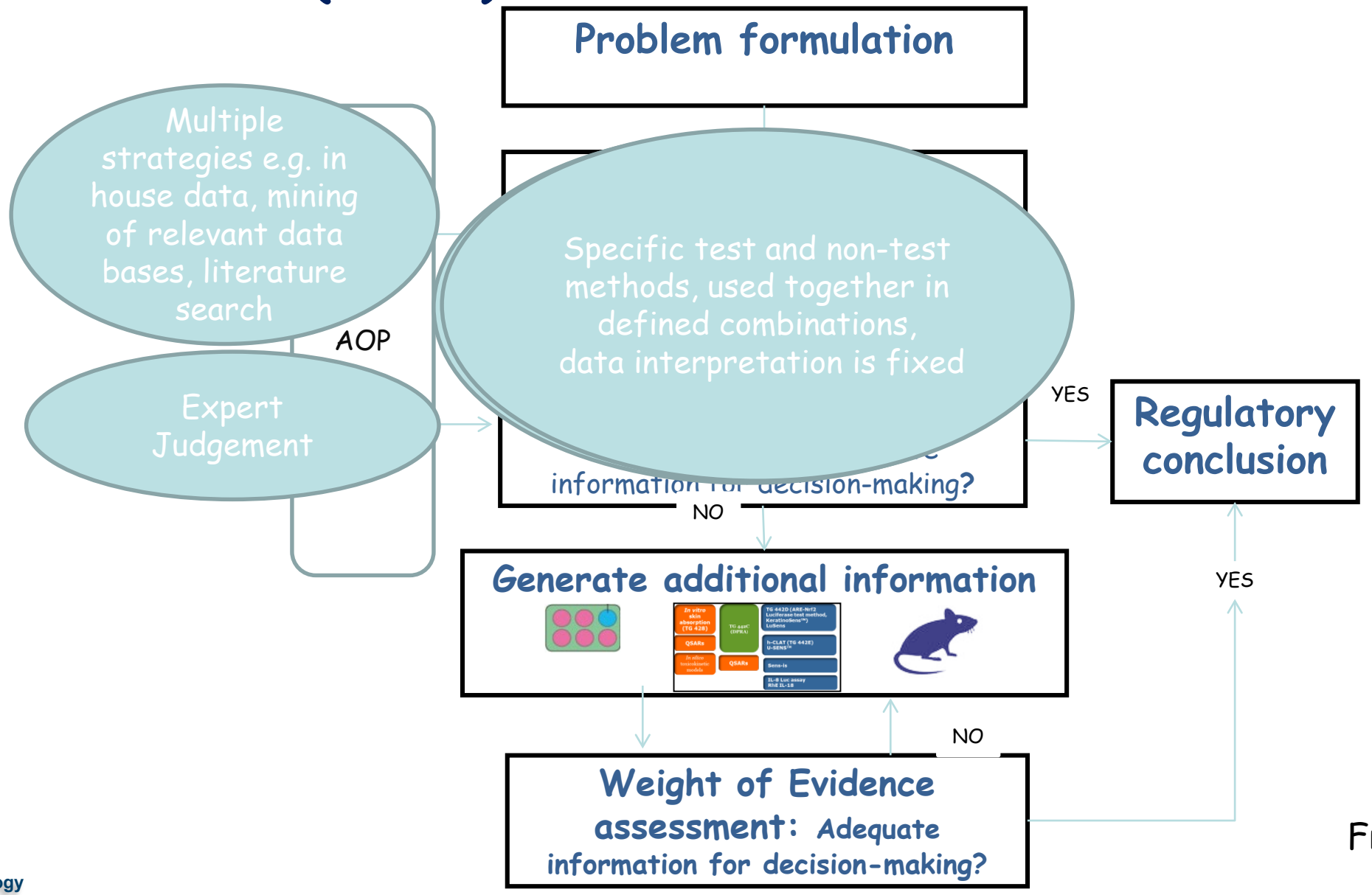
Typical Information within an IATA: IATA elements

- Historical information on the chemical of interest
- Non-standard *in vivo* tests
- Information from "similar" chemicals
- Predictions from other 'non-testing' approaches such as (Q)SAR
- *In chemico* tests
- *In vitro* tests
- Molecular biology, -omics
- Exposure, (bio-)kinetics

Mechanistic based and AOP-informed IATA

- As noted, there is a shift towards non animal alternatives as a response to regulatory drivers
- Integration of different non-animal approaches requires an organising framework to ensure that the different information sources are being interpreted in their appropriate context. This is particularly relevant for New Approach Methodologies (NAMs).
- AOPs serve to provide this organisational framework and hence play an important role in developing and applying IATA for different purposes as well as provide a roadmap for future QSAR development
- AOPs provide the linkage from chemistry, through the Molecular Initiating Event (MIE) to Adverse Effect
- Data from key events provides support to, and will enhance, read-across especially for regulatory acceptance as well as supports definition of domains for MIEs

General workflow in Integrated Approaches to Testing and Assessment (IATA)



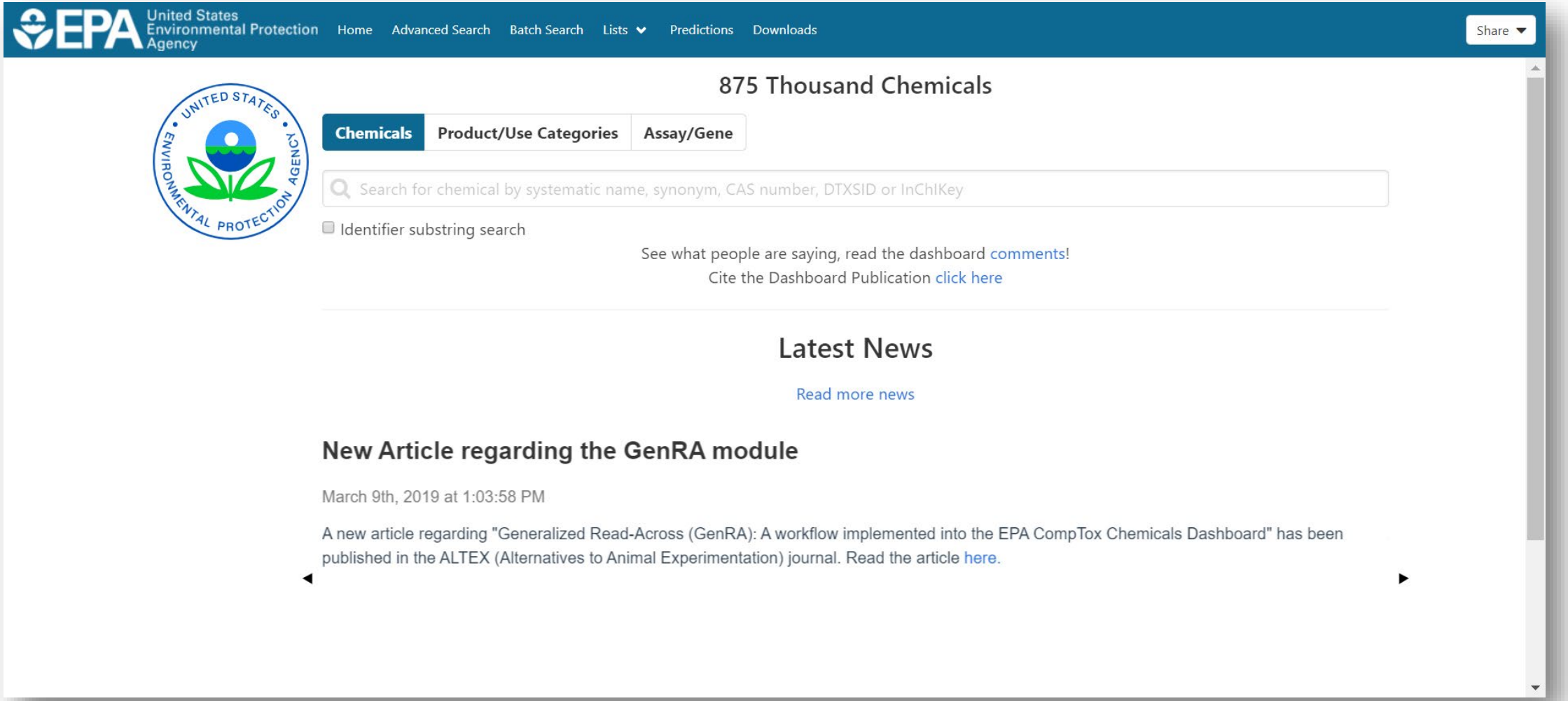
From OECD

EPA CompTox Chemicals Dashboard

- A publicly accessible website delivering access:
  - ~875,000 chemicals with related property data
  - Experimental and predicted physicochemical property data
  - Integration to "biological assay data" for 1000s of chemicals
  - Information regarding consumer products containing chemicals
  - Links to other agency websites and public data resources
  - "Literature" searches for chemicals using public resources
  - "Batch searching" for thousands of chemicals
  - DOWNLOADABLE Open Data for reuse and repurposing

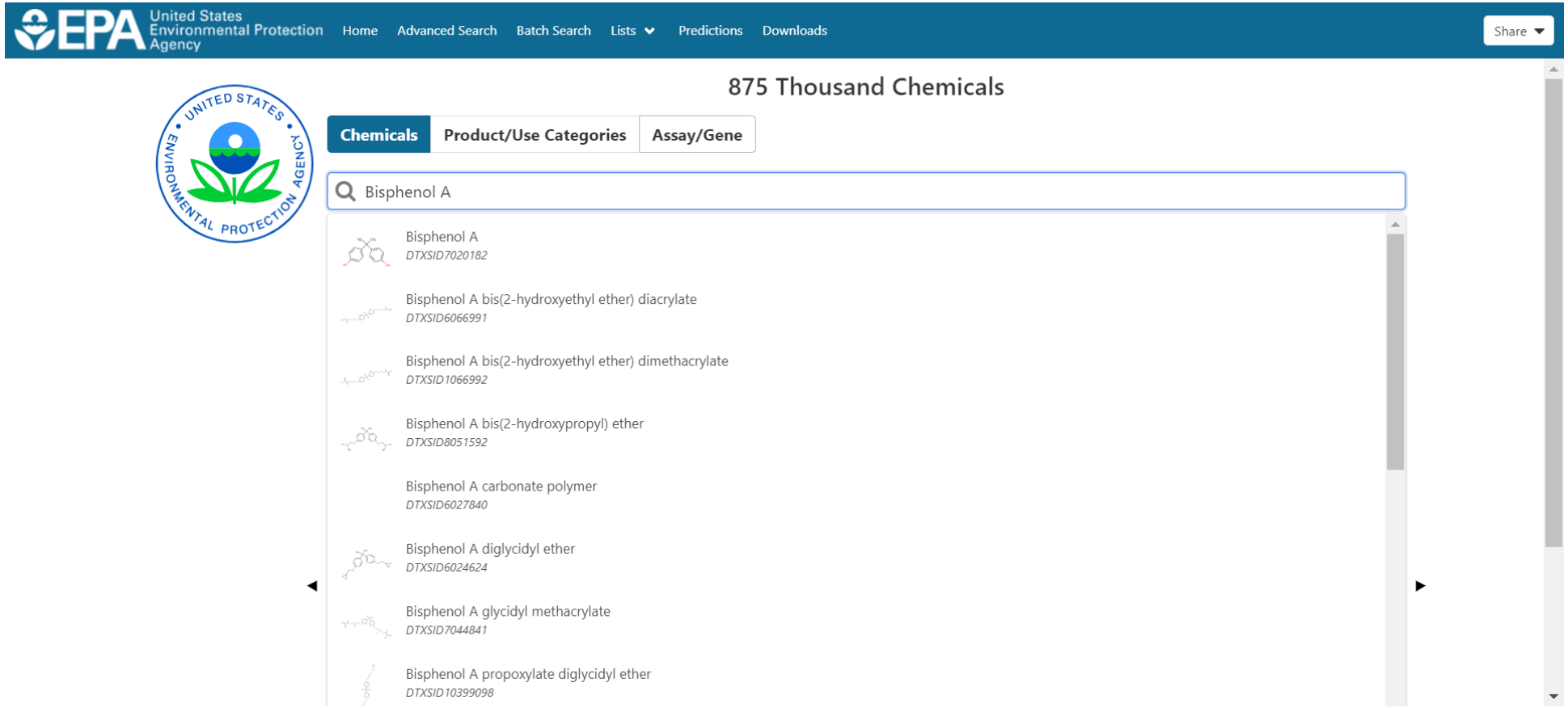
https://comptox.epa.gov/

CompTox Chemicals Dashboard: Landing Page

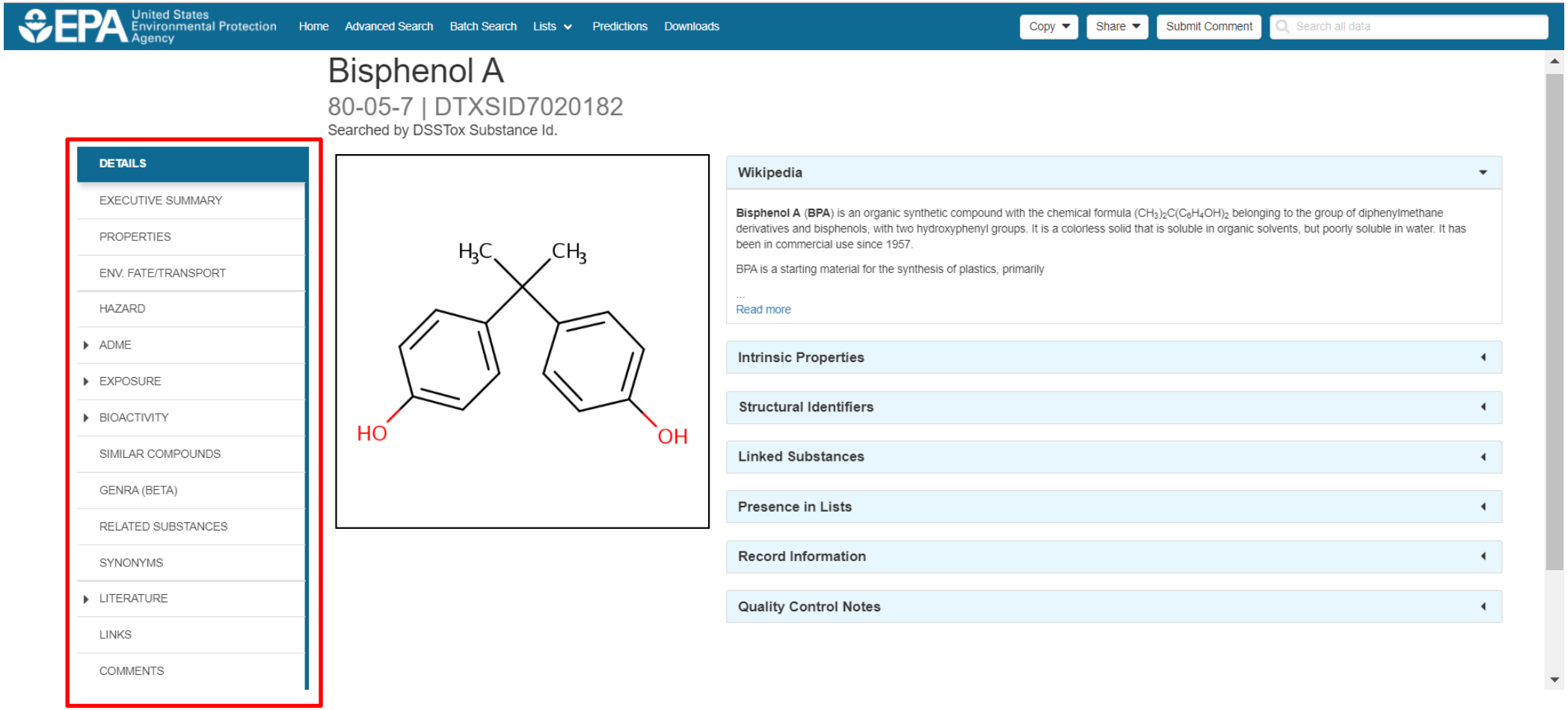


CompTox Chemicals Dashboard: Landing Page

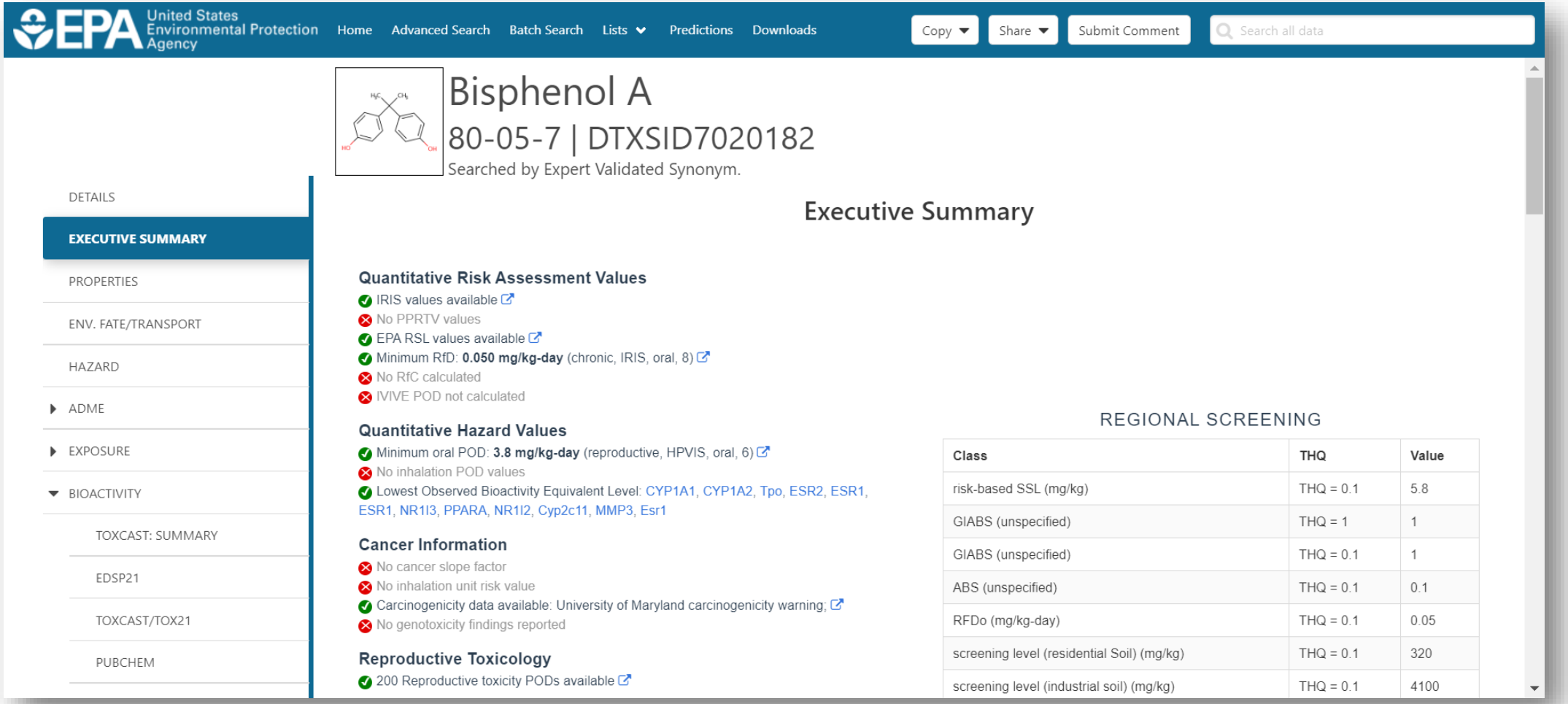
- Different entry points depending on domain of interest



CompTox Chemicals Dashboard: Landing Page for a specific chemical



CompTox Chemicals Dashboard: Executive Summary



Computational toxicology tools add value to most regulatory decisions

- Prioritisation
- Screening level hazard assessment
- Risk Assessment
- Exposure Assessment

Risk-Based prioritisation

- Could involve a combination of available experimental data and new approach methods (NAMs) such as HTTR, HTS
- One approach considered involved coupling Threshold of Toxicological Concern (TTC) with High Throughput Exposure (HTE) modelling to rank order substances for further evaluation
- TTC is a principle that refers to the establishment of a human exposure threshold value for (groups of) chemicals below which there would be no appreciable risk to human health
- Relies on past accumulated knowledge regarding the distribution of potencies of relevant classes of chemicals for which good toxicity data do exist

TTC is based on a predicted tumour risk of 1 in a million, derived through an analysis of genotoxic chemicals

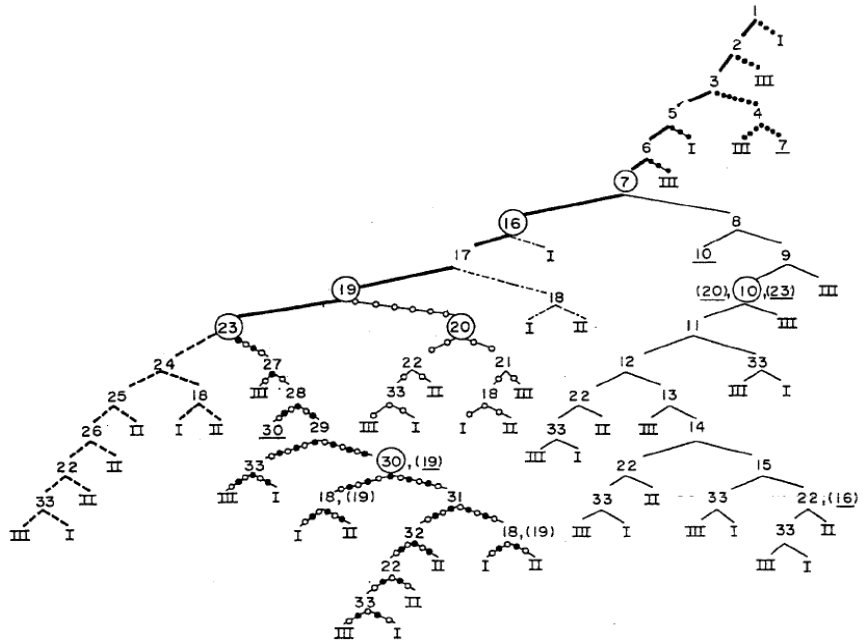
TTC is based on frequency distributions (5th percentile) of NO(A)ELs of non-genotoxic chemicals

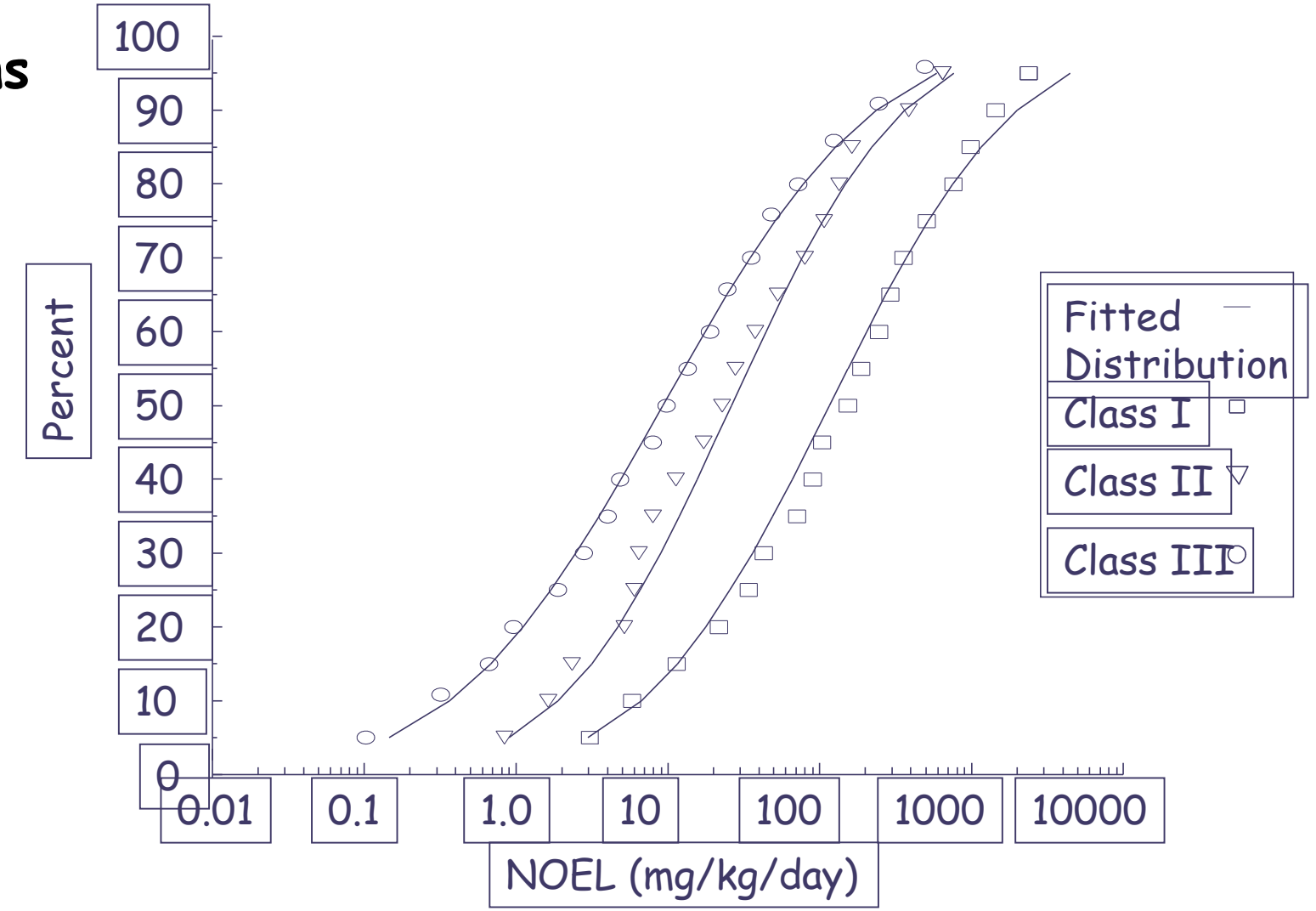
TTC values

Type of substance	µg/person/day (µg/kg-day for 60 kg adult)
Alerts for potential genotoxic carcinogenicity	Kroes: 0.15 (0.0025 µg/kg-day) ICH: 1.5 (0.025 µg/kg-day)
Acetylcholinesterase inhibitors (AChEI) Organophosphate/carbamate	18 (0.3 µg/kg-day)
Cramer Class III	90 (1.5 µg/kg-day)
Cramer Class II	540 (9.0 µg/kg-day)
Cramer Class I	1800 (30 µg/kg-day)

Cumulative Distributions of Cramer Structural Class NOELs

- Decision tree of 33 questions



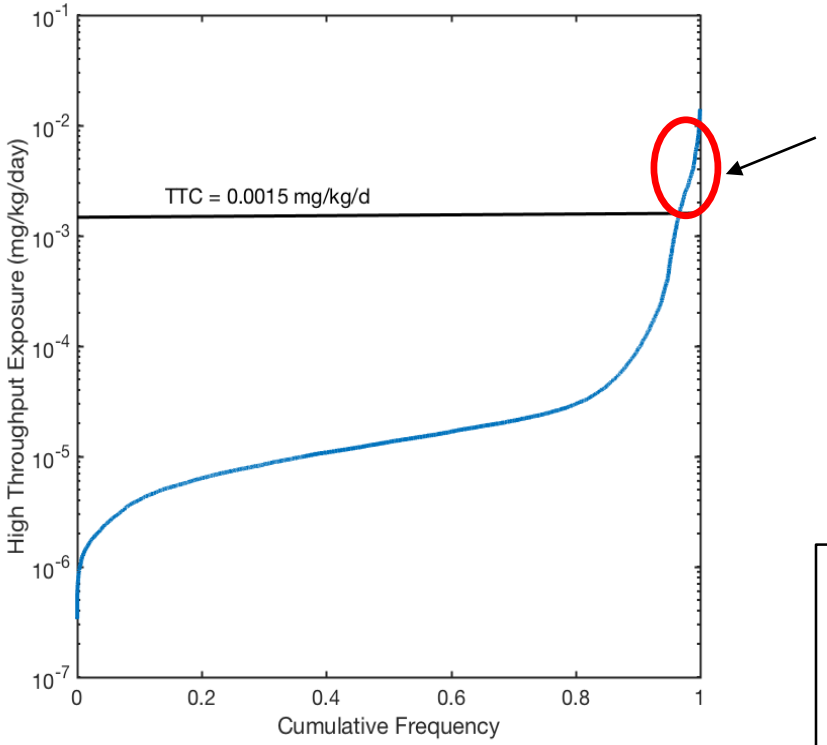


Predicted HT exposures

- Wambaugh and colleagues (2014) developed a rapid heuristic high throughput exposure (HTE) model that enables prediction of potential human exposure to thousands of substances for which little or no empirical exposure data are available.
- The HTE model was calibrated by comparison to NHANES urinary data that reflects total exposure (all routes/sources)

Integrating TTC with predicted HT exposures

- Compared the conservative Cramer Class III TTC value of 1.5 µg/kg-day to the previously calculated median and upper 95% credible interval (UCI) of total daily median exposure rates for 7968 chemicals

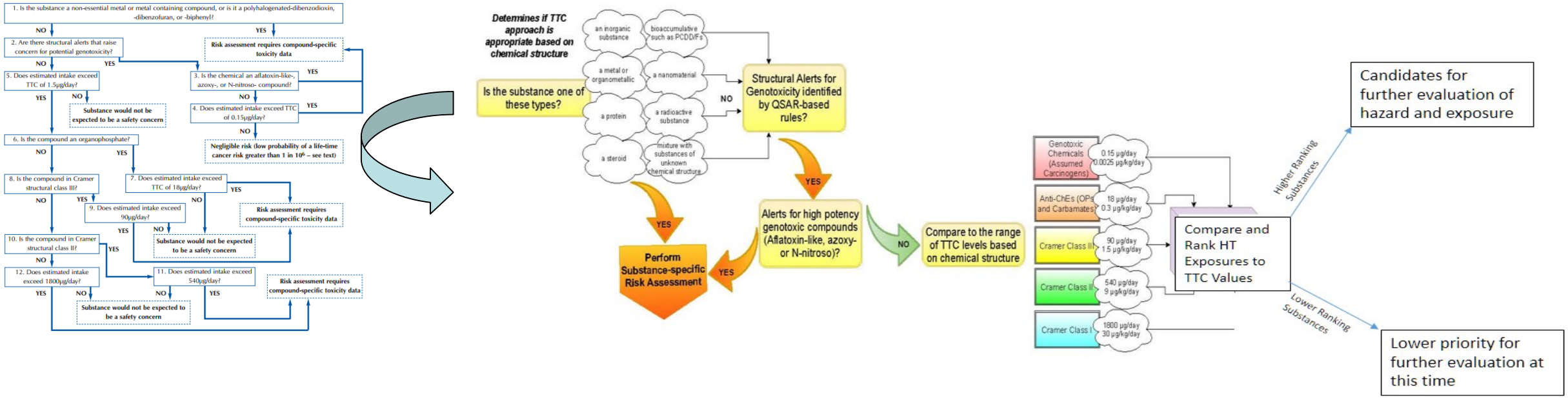


only 273 (fewer than 5%) were found to have UCI daily exposures estimates that exceeded the Cramer Class III TTC value of 1.5 µg/kg-day

Initial evaluation showed the approach of using the ratio of exposure to TTC (HTE: TTC) appeared promising for risk-based prioritisation

Risk-Based prioritisation

- Refined the approach using the Kroes et al structure-based workflow for TTC



- None of the substances categorised as Cramer Class I or Cramer Class II exceeded their respective TTC values.
- No more than 2% of substances categorised as Cramer Class III or acetylcholinesterase inhibitors exceeded their respective TTC values.
- Majority of chemicals with genotoxicity structural alerts did exceed the relevant TTC – recommendations were proposed for next steps

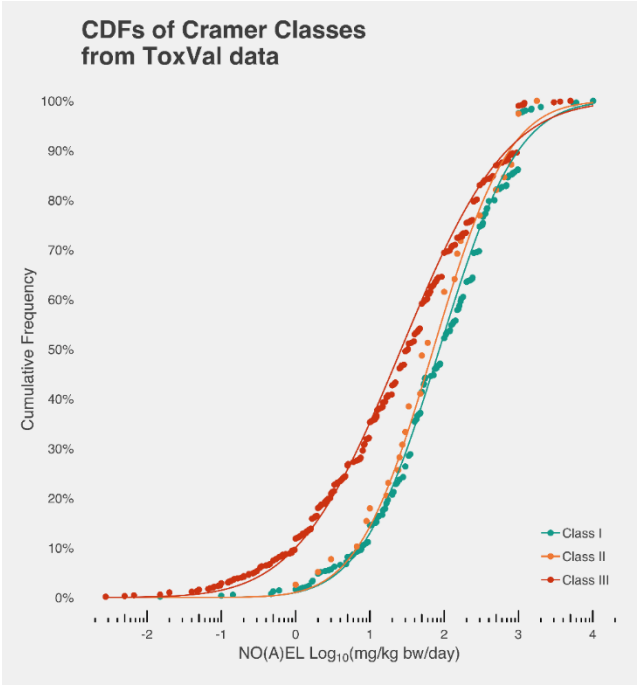
Patlewicz et al, 2018

Risk-Based prioritisation

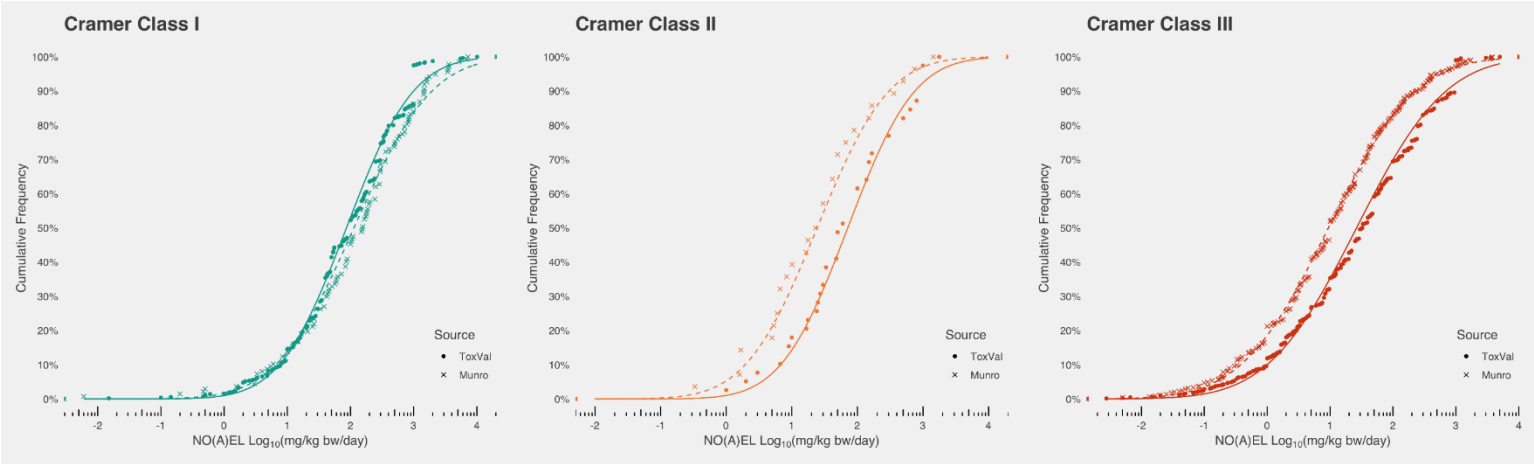
- Investigate relevance of existing TTC values for substances of interest to EPA
- Extracted data from EPA's ToxValDB, which aggregates *in vivo* testing data from over 40 sources including US federal and state agencies, as well as international agencies such as the European Chemicals Agency and the World Health Organisation
- Objectives were:
  - Reproduce the TTC values developed by Munro et al (1996)
  - Follow the Kroes et al (2004) workflow to assign substances present in ToxVal to their respective Cramer classes and use the associated repeat dose toxicity data to derive new TTC values
  - Evaluate whether the TTC values from ToxVal and Munro are statistically equivalent
  - Derive confidence intervals for the new TTC values
  - Compare and contrast the chemistry of the two data sets to rationalise any (dis)similarities in TTC values

Risk-Based prioritisation

Follow the Kroes et al (2004) workflow to assign substances present in ToxVal to their respective Cramer classes and use the associated repeat dose toxicity data to derive new TTC values

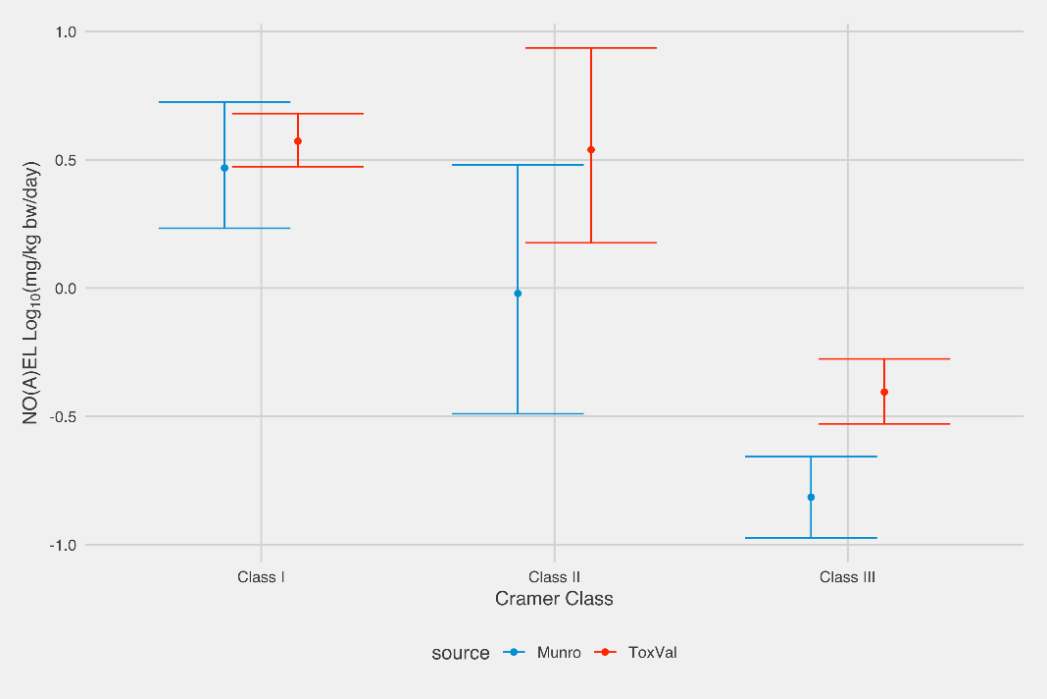


Evaluate whether the TTC values from ToxVal and Munro are statistically equivalent & derive confidence intervals for the new TTC values



Risk-Based prioritisation

- Bootstrap sampling used to quantify the uncertainty around the 5th percentiles values for both ToxVal and Munro data sets
- Differences were observed for substances assigned as Cramer Class III



- Presence of OP/carbamates in the Munro Cramer class III set largely explained the difference in 5th percentile values

Nelms et al, submitted

Definitions: Chemical grouping approaches

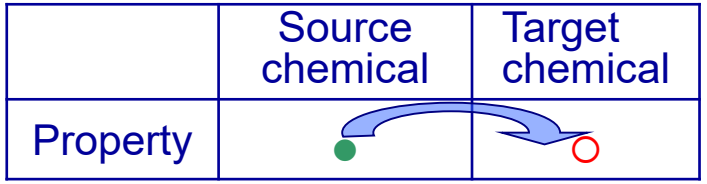
"Analogue approach" refers to grouping based on a very limited number of chemicals (e.g. target substance + source substance)

"Category approach" is used when grouping is based on a more extensive range of analogues (e.g. 3 or more members)

A chemical category is a group of chemicals whose physico-chemical and human health and/or environmental toxicological and/or environmental fate properties are likely to be similar or follow a regular pattern as a result of structural similarity (or other similarity characteristics).

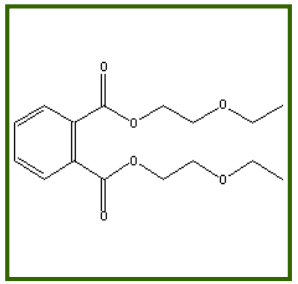
Definitions: Read-across

- Read-across describes the method of filling a data gap whereby a chemical with existing data values is used to make a prediction for a 'similar' chemical.
- A target chemical is a chemical which has a data gap that needs to be filled i.e. the subject of the read-across.
- A source analogue is a chemical that has been identified as an appropriate chemical for use in a read-across based on similarity to the target chemical and existence of relevant data.



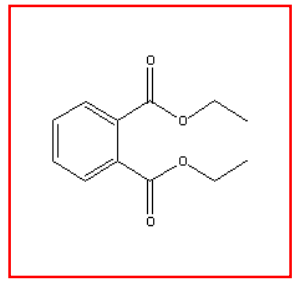
● Reliable data

○ Missing data



Known to be harmful

Acute toxicity?



Predicted to be harmful

A harmonised hybrid read-across workflow

Contents lists available at ScienceDirect

Computational Toxicology

journal homepage: www.elsevier.com

Journal Cover Image

Navigating through the minefield of read-across frameworks: A commentary perspective

Grace Patlewicz[a, *], Mark T.D. Cronin[b], George Helman[a, c], Jason C. Lambert[d], Lucina E. Lizarraga[d], Imran Shah[a]

[a] National Center for Computational Toxicology (NCCT), Office of Research and Development, US Environmental Protection Agency (US EPA), 109 TW Alexander Dr, Research Triangle Park (RTP), NC 27711, USA
[b] School of Pharmacy and Biomolecular Sciences, Liverpool John Moores University, Byrom Street, Liverpool L3 3AF, UK
[c] Oak Ridge Institute for Science and Education (ORISE), 1299 Bethel Valley Road, Oak Ridge, TN 37830, USA
[d] National Center for Evaluation Assessment (NCEA), US Environmental Protection Agency (US EPA), 26 West Martin Luther King Dr, Cincinnati, OH 45268, USA

A harmonised hybrid read-across workflow

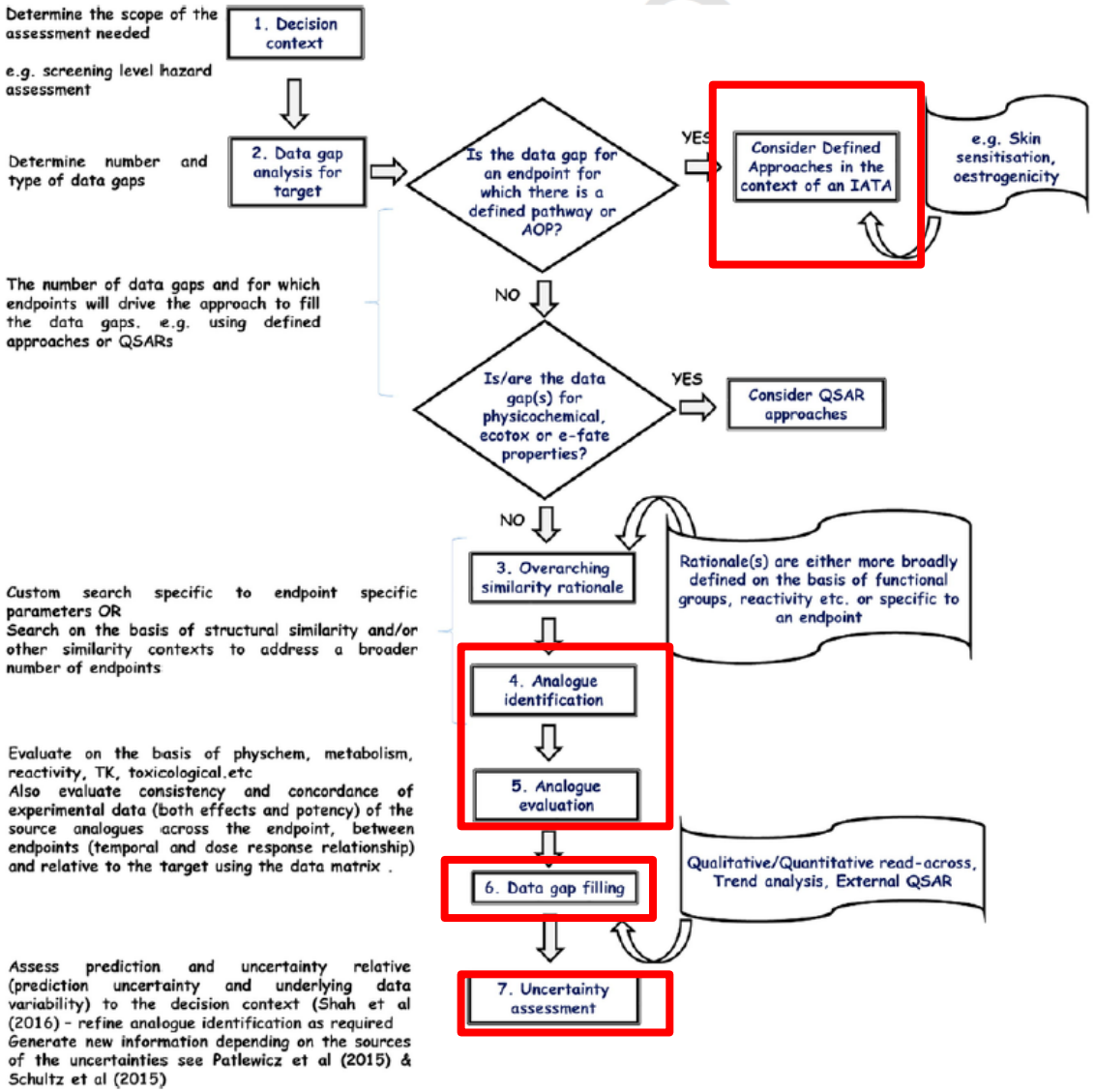


Fig. 9. A harmonised hybrid development and assessment framework.

Where do other NAM fit?
How should we transition to data-driven approaches?
What about characterising the uncertainty of the predictions made?

Patlewicz et al., 2018

Selected read-across tools

Computational Toxicology 3 (2017) 1–18

Contents lists available at ScienceDirect

Computational Toxicology

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

Navigating through the minefield of read-across tools: A review of in silico tools for grouping

Grace Patlewicz [a,*], George Helman [a,b], Prachi Pradeep [a,b], Imran Shah [a]

[a] National Center for Computational Toxicology (NCCT), Office of Research and Development, US Environmental Protection Agency, 109 TW Alexander Dr, Research Triangle Park (RTP), NC 27711, USA
[b] Oak Ridge Institute for Science and Education (ORISE), Oak Ridge, TN, USA

ARTICLE INFO

Article history:
Received 29 March 2017
Received in revised form 22 May 2017
Accepted 25 May 2017
Available online 29 May 2017

Keywords:
Category approach
Analogue approach
Data gap filling
Read-across
(Q)SAR
Trend analysis
Nearest neighb

ABSTRACT

Read-across is a popular data gap filling technique used within analogue and category approaches for regulatory purposes. In recent years there have been many efforts focused on the challenges involved in read-across development, its scientific justification and documentation. Tools have also been developed to facilitate read-across development and application. Here, we describe a number of publicly available read-across tools in the context of the category/analogue workflow and review their respective capabilities, strengths and weaknesses. No single tool addresses all aspects of the workflow. We highlight how the different tools complement each other and some of the opportunities for their further development to address the continued evolution of read-across.

Published by Elsevier B.V.

Selected read-across tools

Tool	AIM	ToxMatch	AMBIT	OECD Toolbox	CBRA	ToxRead	GenRA
Analogue identification	X	X	X	X	X	X	X
Analogue Evaluation	NA	X	X by other tools availabl e	X	X	X For Ames & BCF	NA
Data gap analysis	NA	X	X Data matrix can be exporte d	X Data matrix viewable	NA	NA	X Data matrix can be exported
Data gap filling	NA	X	User driven	X	X	X	X
Uncertainty assessment	NA	NA	NA	X	NA	NA	X
Availability	Free	Free	Free	Free	Free	Free	Free

GenRA (Generalised Read-Across)

- Predicting toxicity as a similarity-weighted activity of nearest neighbours based on chemistry and bioactivity descriptors (Shah et al, 2016)
- Generalised version of the Chemical-Biological Read-Across (CBRA) developed by Low et al (2013)
- Goal: To establish an objective performance baseline for read-across and quantify the uncertainty in the predictions made

$$y_i^{\beta,\alpha} = \frac{\sum_j^k s_{ij}^{\alpha} x_j^{\beta}}{\sum_j^k s_{ij}^{\alpha}}$$

Jaccard similarity:

$$s_{ij} = \frac{\sum_l (x_{il} \wedge x_{jl})}{\sum_l (x_{il} \vee x_{jl})}$$

$\alpha \in \{chm, bio, bc\}$

$\beta \in \{bio, tox\}$

y_i = predicted activity of chemical (c_i)

x_j^{β} = activity of c_j in β

s_{ij}^{α} = Jacccard similarity between $\boldsymbol{x}_i^{\alpha}, \boldsymbol{x}_j^{\alpha}$

k = up to k nearest neighbours

GenRA v1.0 - Approach

I. Data

1,778 Chemicals
3,239 Structure descriptors (chm)
820 Bioactivity hitcall (bio)
ToxCast

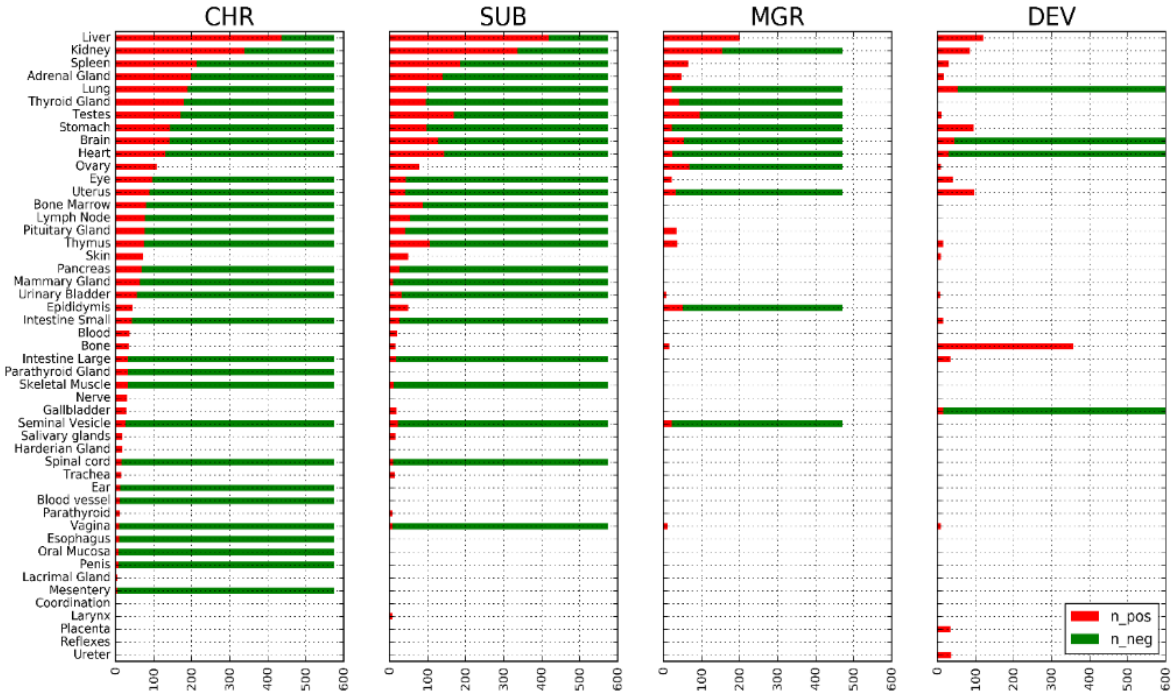
574 toxicity effects (tox)
ToxRefDB

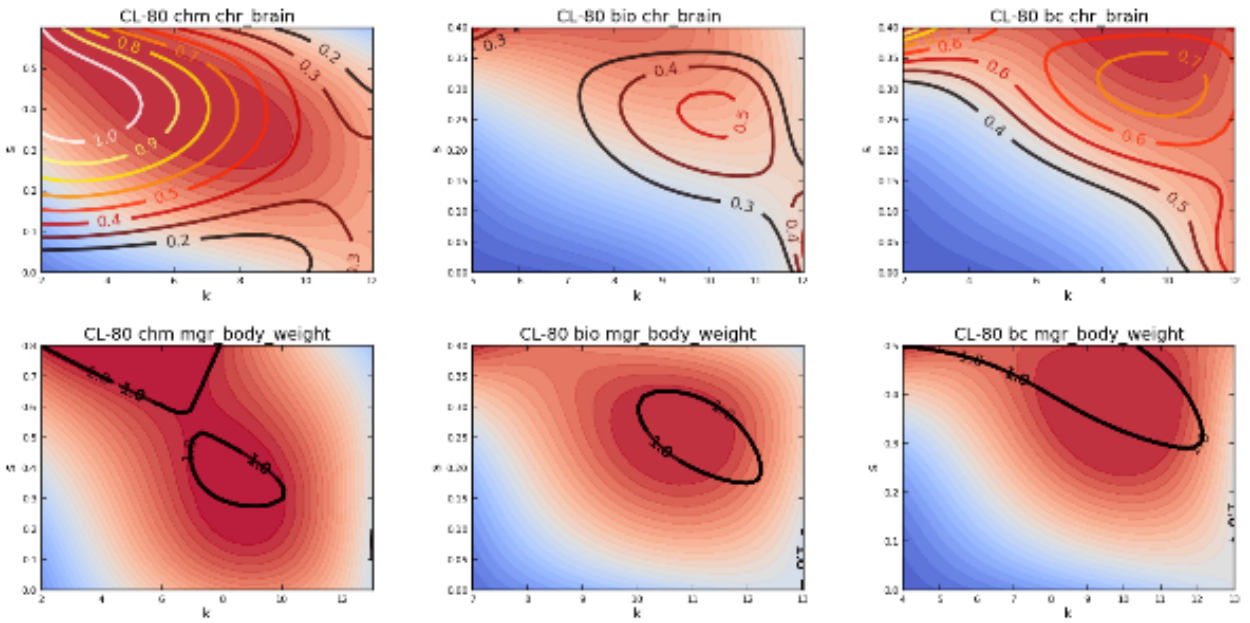
II. Define Local neighbourhoods

Use K-means analysis to group chemicals by similarity
Use cluster stability analysis
~ 100 local neighbourhoods

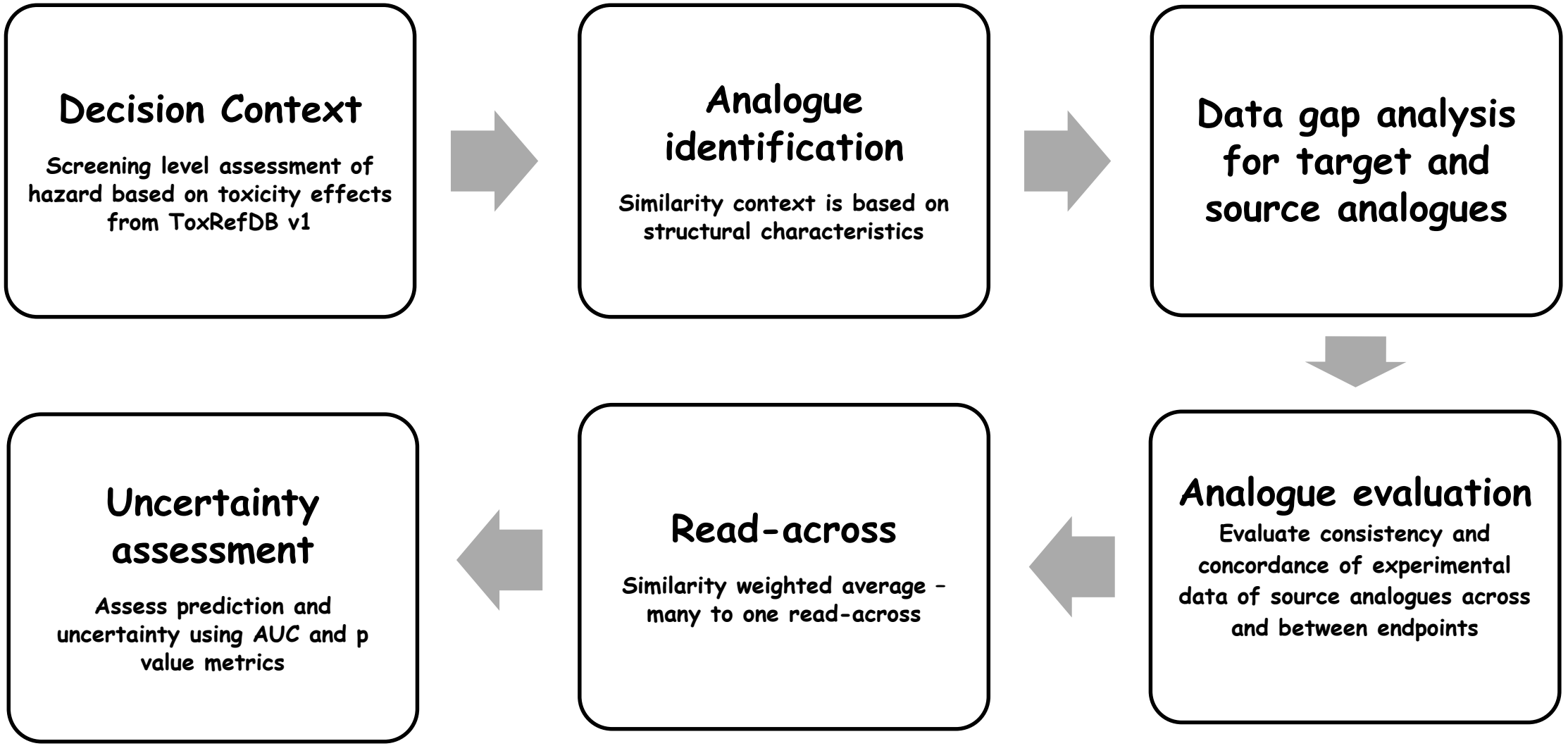
III. GenRA

Use GenRA to predict toxicity effects in local neighbourhoods
Evaluate impact of structural and/or bioactivity descriptors on prediction
Quantify uncertainty



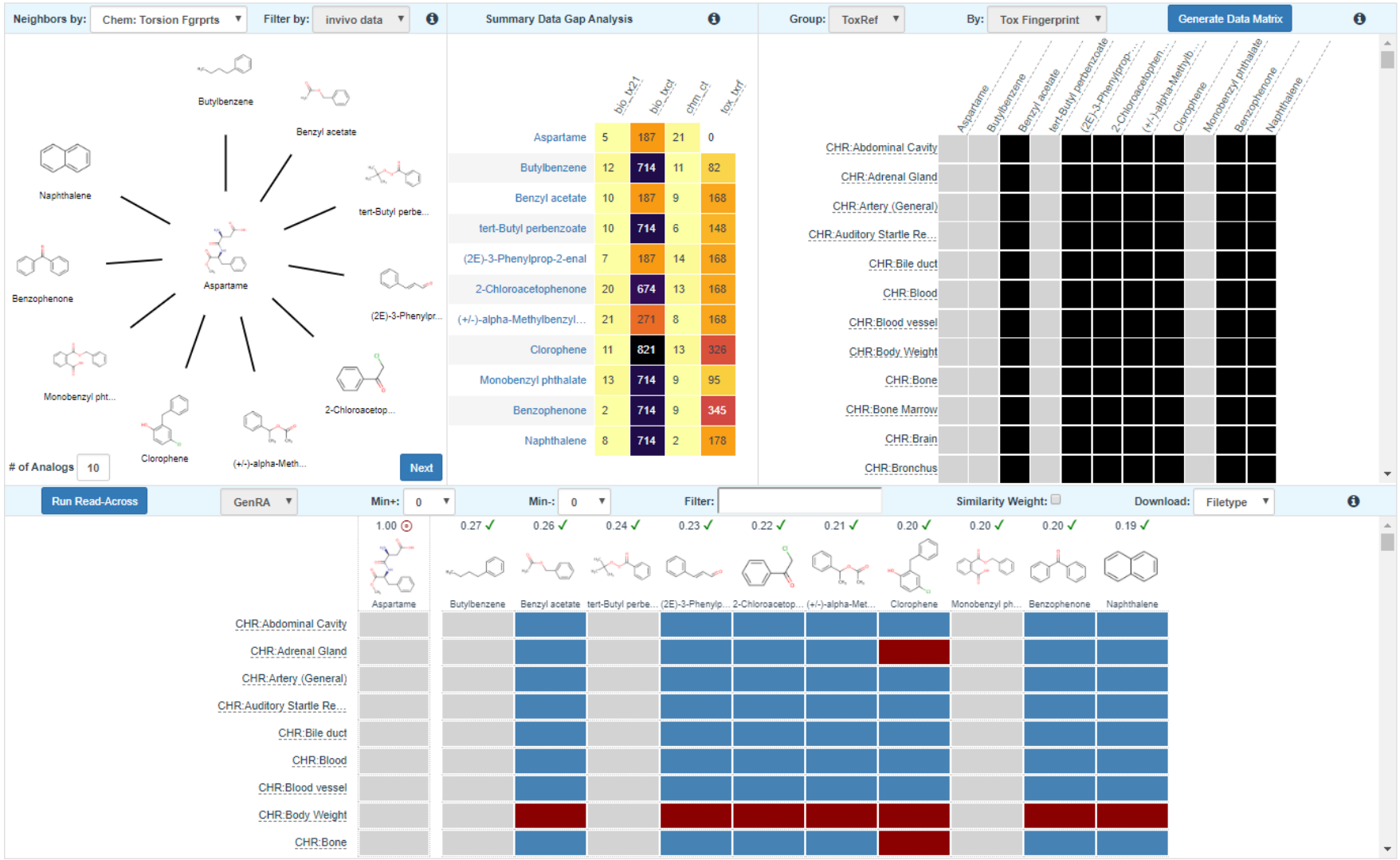


Read-across workflow in GenRA v1.0



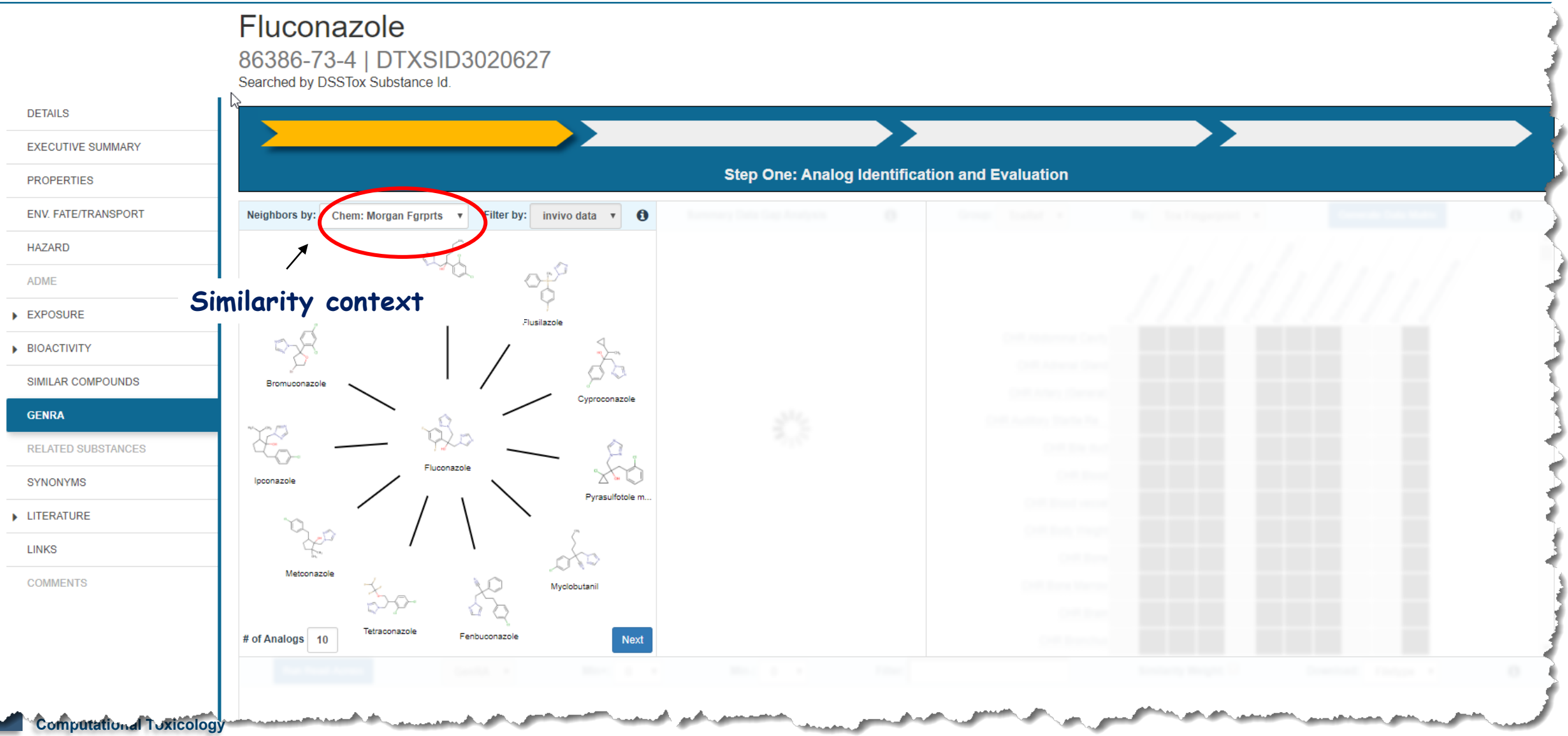
GenRA tool in reality

- Integrated into the EPA CompTox Chemicals dashboard



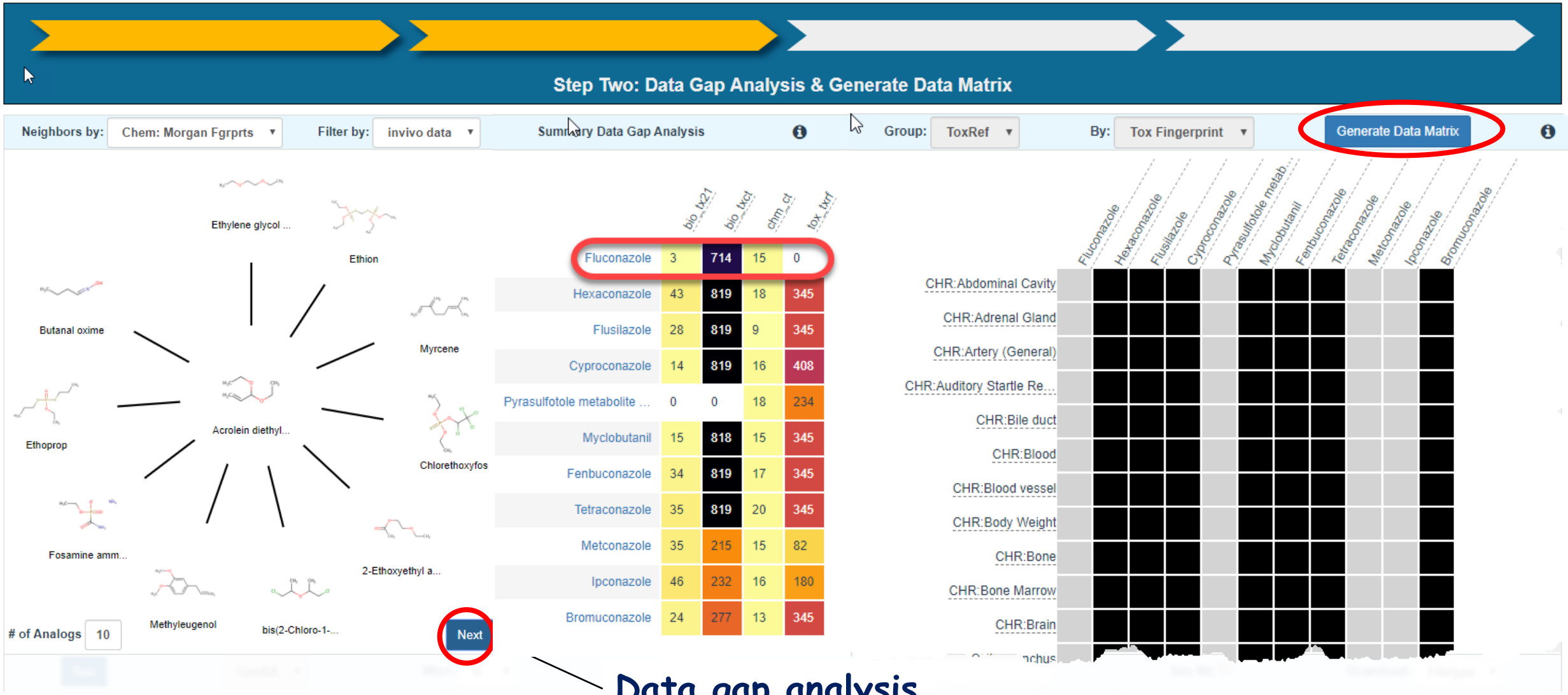
GenRA tool in practice

- Structured as a workflow



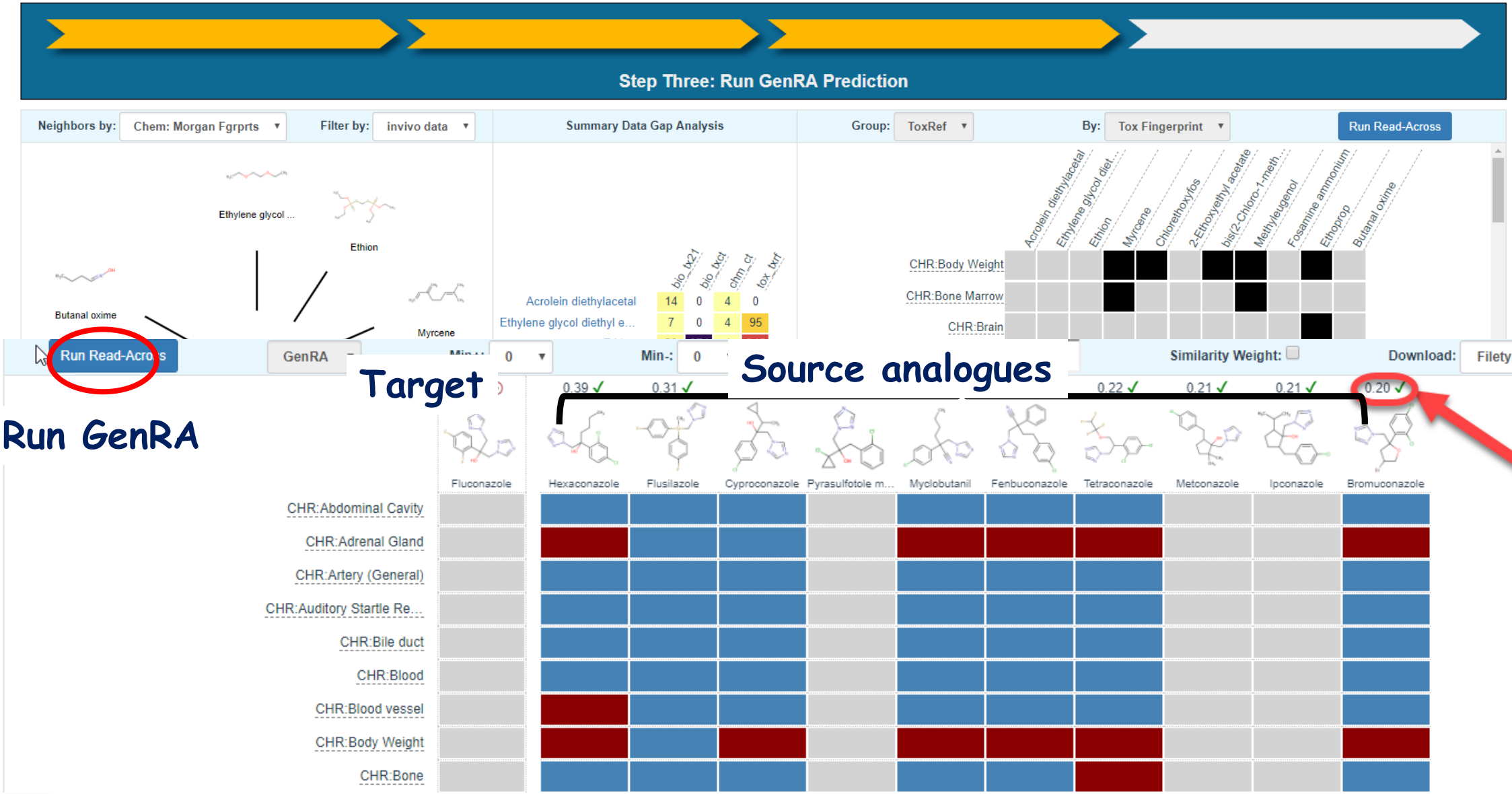
GenRA tool in practice

GenRA

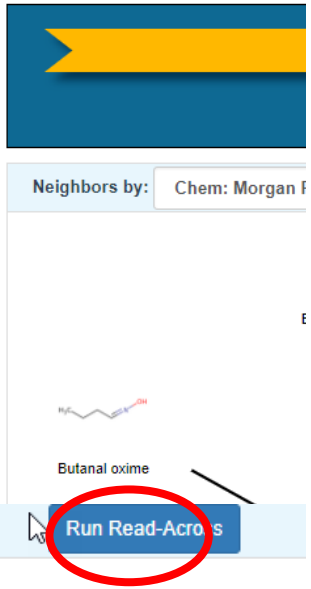


Data gap analysis

GenRA tool in practice



GenRA tool in practice



Run GenRA

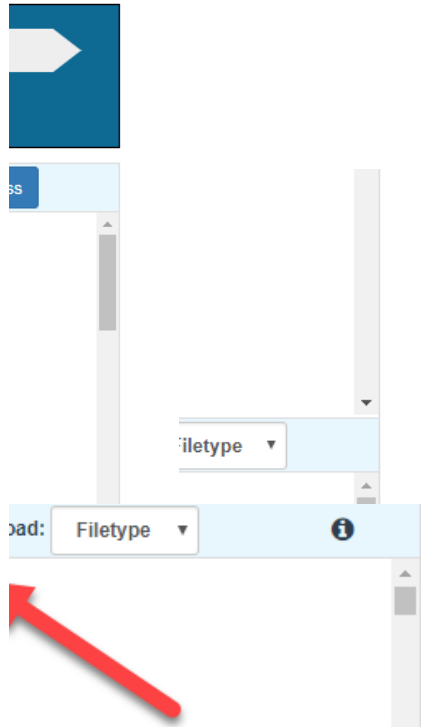
ALTEX preprint
published February 4, 2019
doi:10.14573/altex.1811292

Short Communication

Generalized Read-Across (GenRA): A workflow implemented into the EPA CompTox Chemicals Dashboard

George Helman[1,2], Imran Shah[2], Antony J. Williams[2], Jeff Edwards[2], Jeremy Dunne[2] and Grace Patlewicz[2]*

[1]Oak Ridge Institute for Science and Education (ORISE), Oak Ridge, TN, USA; [2]National Center for Computational Toxicology (NCCT), Office of Research and Development, US Environmental Protection Agency, Research Triangle Park (RTP), NC, USA



Abstract

Generalized Read-Across (GenRA) is a data driven approach which makes read-across predictions on the basis of a similarity weighted activity of source analogues (nearest neighbors). GenRA has been described in more detail in the literature (Shah et al., 2016; Helman et al., 2018). Here we present its implementation within the EPA's CompTox Chemicals Dashboard to provide public access to a GenRA module structured as a read-across workflow. GenRA assists researchers in identifying source analogues, evaluating their validity and making predictions of *in vivo* toxicity effects for a target substance. Predictions are presented as binary outcomes reflecting presence or absence of toxicity together with quantitative measures of uncertainty. The approach allows users to identify analogues in different ways, quickly assess the availability of relevant *in vivo* data for those analogues and visualize these in a data matrix to evaluate the consistency and concordance of the available experimental data for those analogues before making a GenRA prediction. Predictions can be exported into a tab-separated value (TSV) or Excel file for additional review and analysis (e.g., doses of analogues associated with production of toxic effects). GenRA offers a new capability of making reproducible read-across predictions in an easy-to use-interface.

GenRA – Next Steps

- Ongoing research:

- Summarising and aggregating the toxicity effect predictions to guide end users – what effect predictions are we most confident about (digesting & interpreting the predictions more efficiently)

- Consideration of other information to define and refine the analogue selection & evaluation – e.g. physicochemical similarity, metabolic similarity, reactivity similarity, bioactivity similarity (transcriptomics similarity)…
  - EPA New Chemical Categories
  - Quantifying the impact of physicochemical similarity on read-across performance (Helman et al., 2018)

GenRA – Next Steps

- Dose response information to refine scope of prediction beyond binary outcomes
  - Transitioning from qualitative to quantitative predictions – how to apply and interpret GenRA in screening level hazard assessment
  - Starting with quantitative data – e.g. acute rat oral toxicity (Helman et al (2019), ToxRefDB v2 (Helman et al (2019)

Case study: Acute toxicity

- Transitioning GenRA to make quantitative predictions
- Investigated extending GenRA using the acute oral rat systemic toxicity data collected as part of the ICCVAM Acute toxicity workgroup
- NICEATM-NCCT effort to collate a large dataset of acute oral toxicity to evaluate the performance of existing predictive models and investigate the feasibility of developing new models

Acute toxicity: Dataset creation

Database Resource	Rows of Data (number of LD50 values)	Unique CAS
ECHA (ChemProp)	5533	2136
JRC AcutoxBase	637	138
NLM HSDB	4082	2238
OECD (eChemPortal)	10206	2314
PAI (NICEATM)	364	293
TEST (NLM ChemIDplus)	13689	13545

Rat oral LD50s:
16,297 chemicals total
34,508 LD50 values

Require unique LD50 values with mg/kg units

15,688 chemicals total
21,200 LD50 values

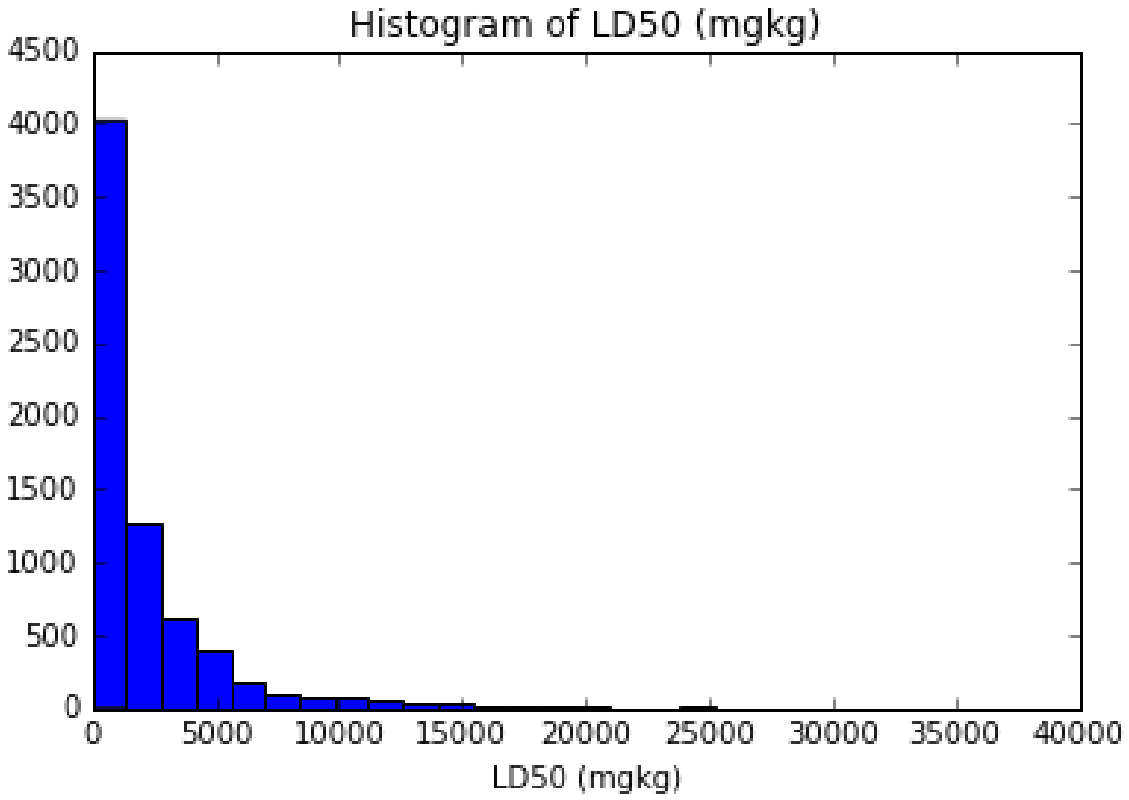
Preprocessing for modelling

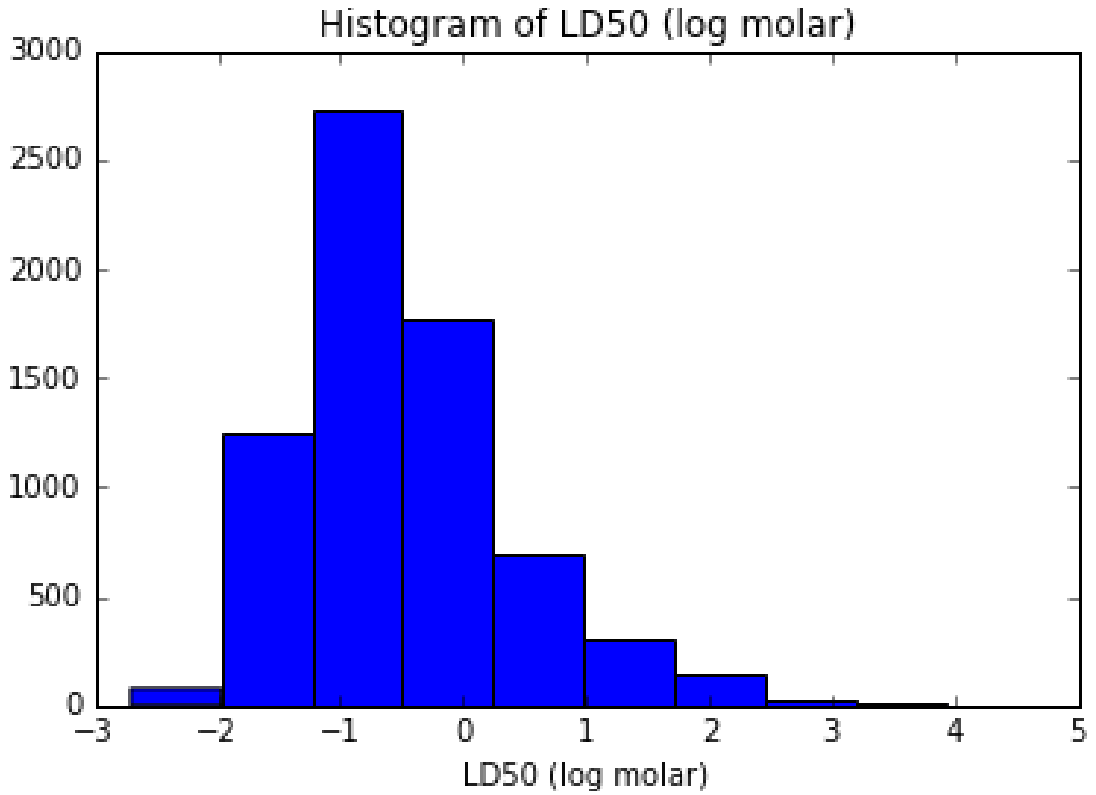
11,992 chemicals
16,173 LD50 values

Karmaus et al, 2018; Kleinstreuer et al., 2018

Exploratory Data Analysis

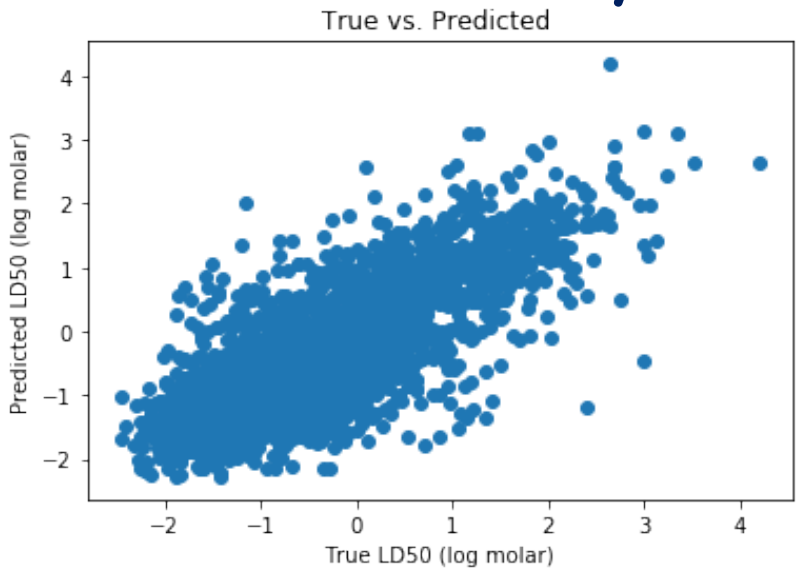
- Found DSSTox matches for 7011 substances
- Extracted MW values

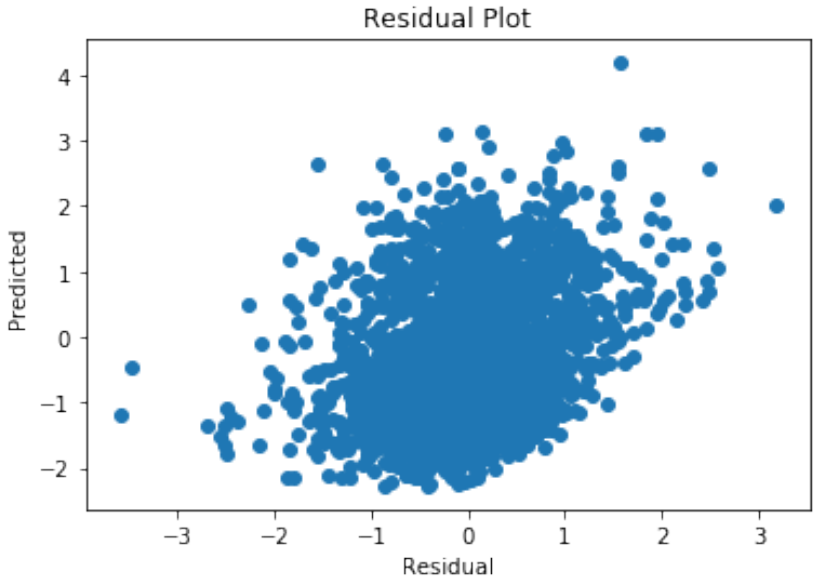




GenRA approach : Overall 'global' performance

- Search for a maximum of 10 nearest neighbours on entire dataset
- Use a min similarity threshold of 0.5

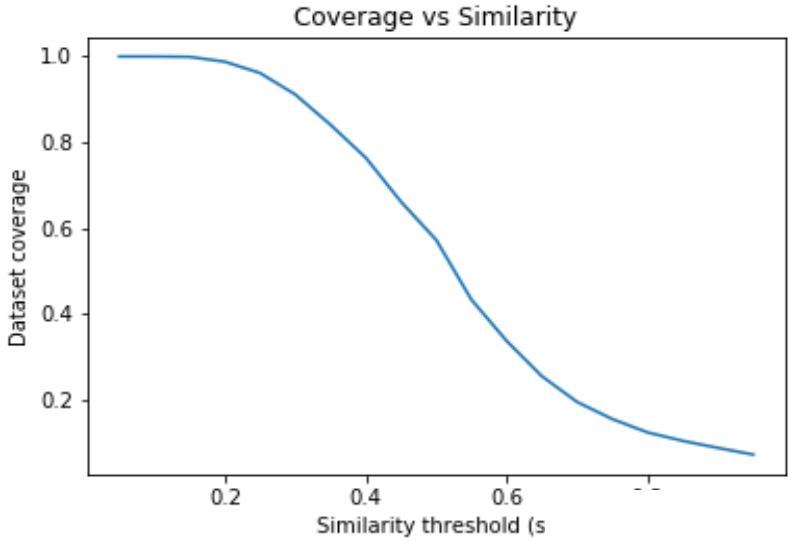




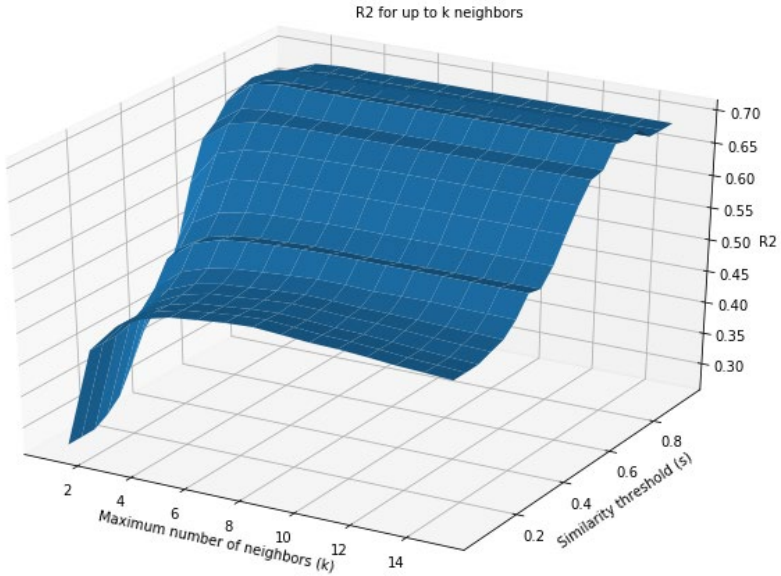
- Linear regression used to fit predicted and observed LD50 values
- R^2 = 0.61
- RMSE = 0.58
- A few outliers, but not too extreme
- Residuals clustered around zero with no obvious patterns

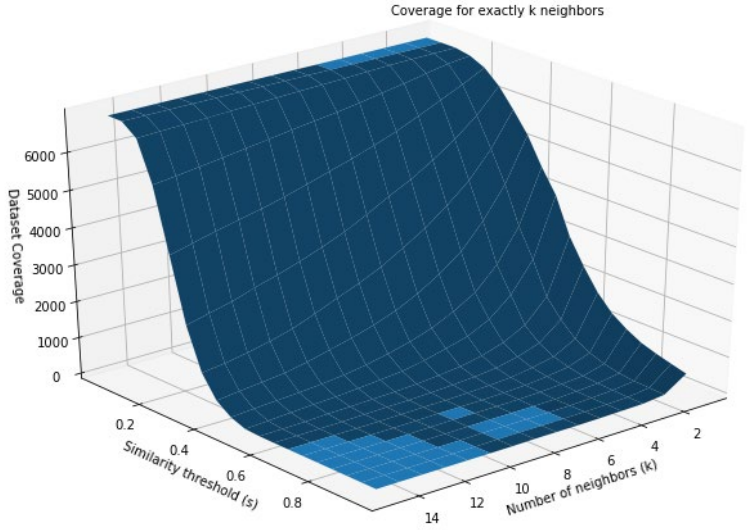
Coverage vs Similarity vs Performance

- Coverage vs Similarity



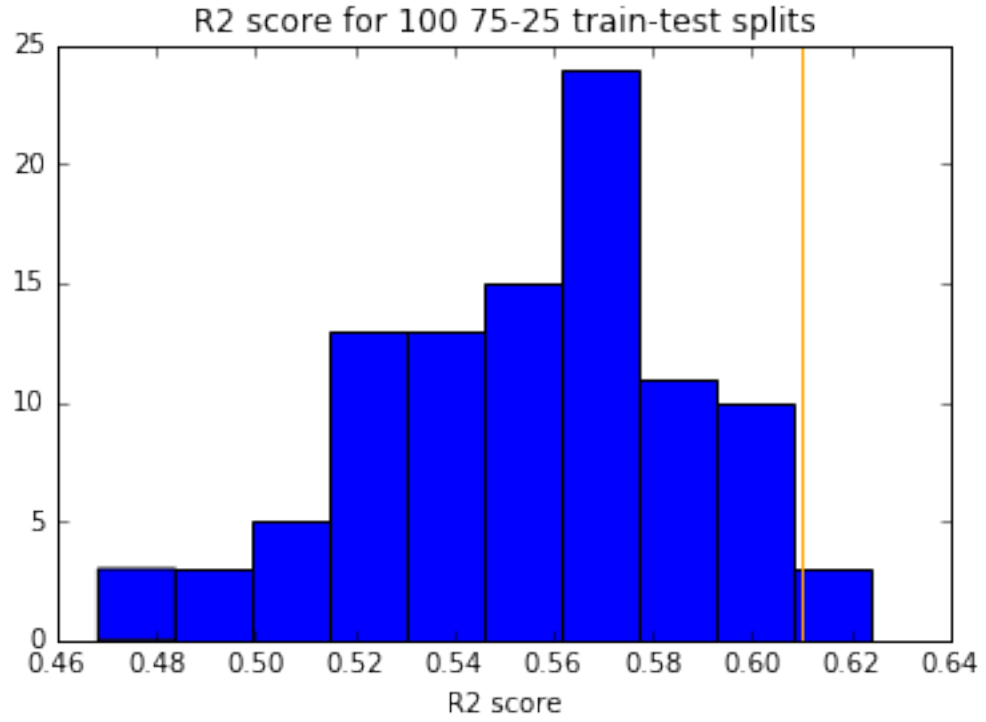
R2 for up to k source analogues





Based on the grid searches performed, k = 10, s = 0.5 were reasonable parameters to tradeoff coverage vs prediction accuracy

Monte Carlo Cross Validation



- Estimate confidence in R2
- 75-25 train-test splits

- R^2 values range from 0.46 to 0.62

- GenRA performs robustly on this acute tox data set

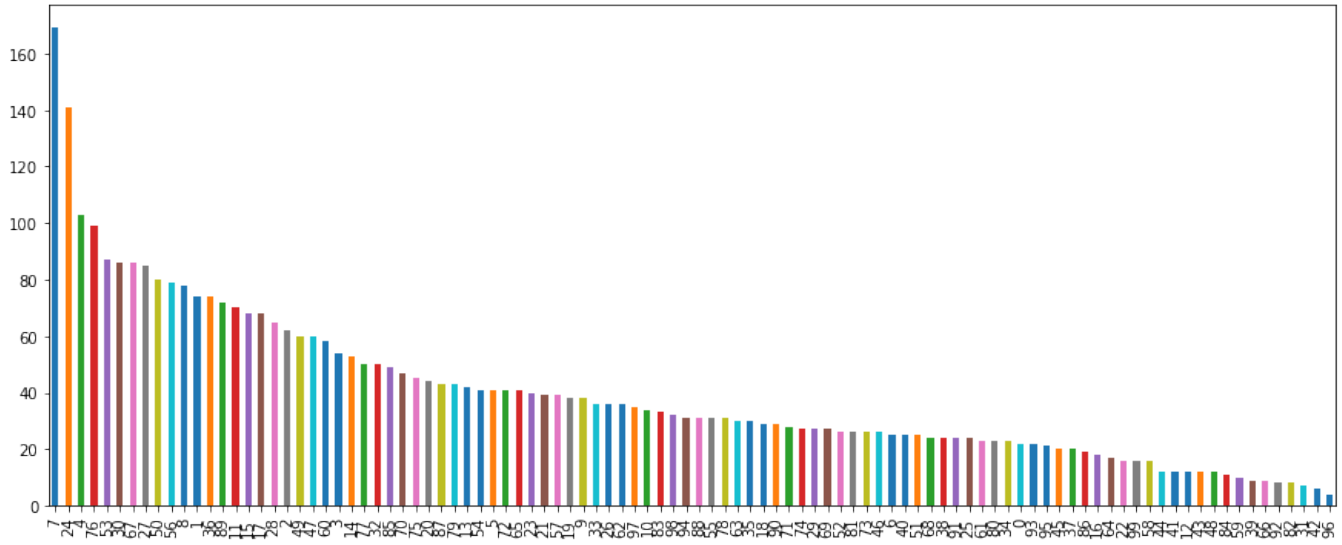
Helman et al. (2019)

Evaluating 'local' performance

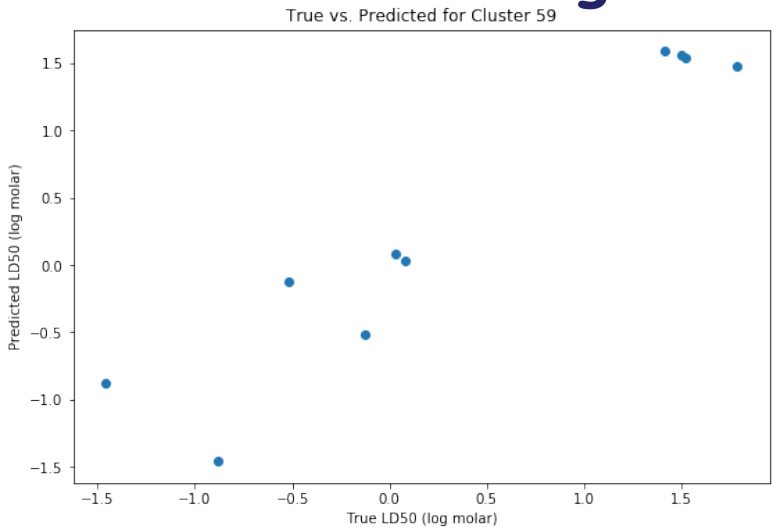
Clustered chemicals into 100 groups on the basis of ToxPrint fingerprints

Explored performance on the basis of individual clusters to gauge what sorts of chemicals resulted in significantly improved performance (R2) relative to the overall 'global' performance reported using 10 nearest neighbours and a similarity of 0.5

Average R2 values improved (R2>0.61) for 19 out of the 100 clusters, some up to 0.91

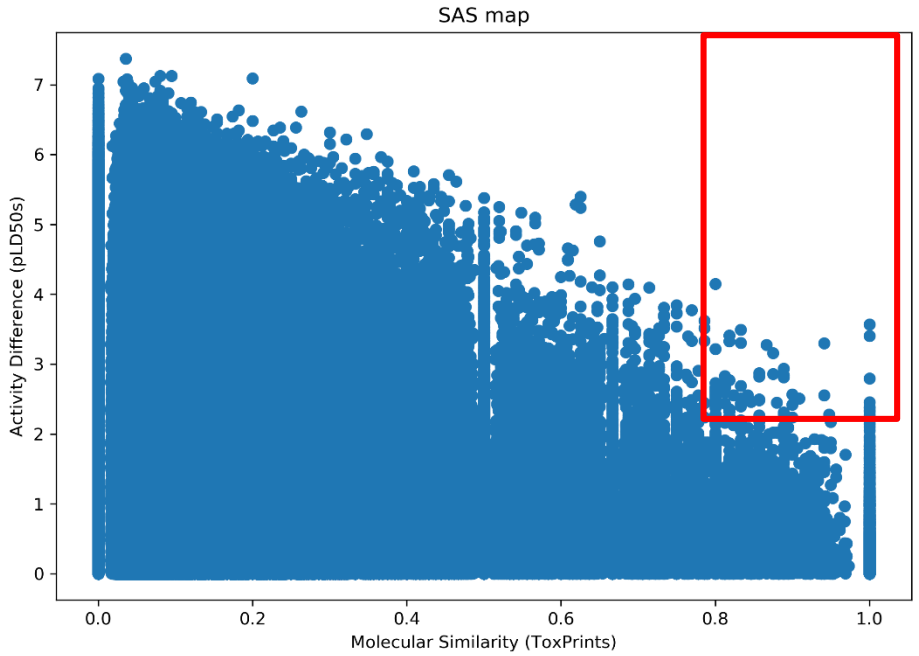


Carbamate containing substances



Structure-Activity similarity (SAS) map

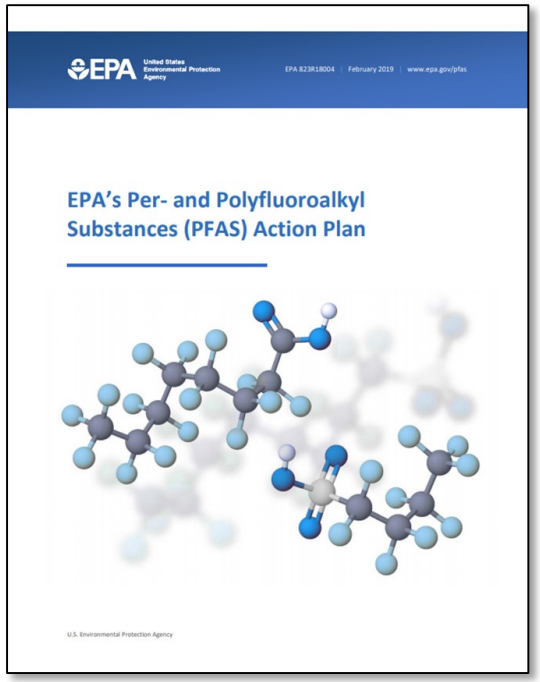
- Are there pairs of substances that are very similar structurally with very high LD50 differences, so called activity cliffs



The number of chemical pairs that fell within the activity cliff quadrant was very low relative to the total number of chemical pairs captured.

This suggests that the chemical fingerprints were able to capture sufficient information to make robust predictions of acute oral toxicity.

EPA Using New Approach Methods to Help Fill Information Gaps for PFAS



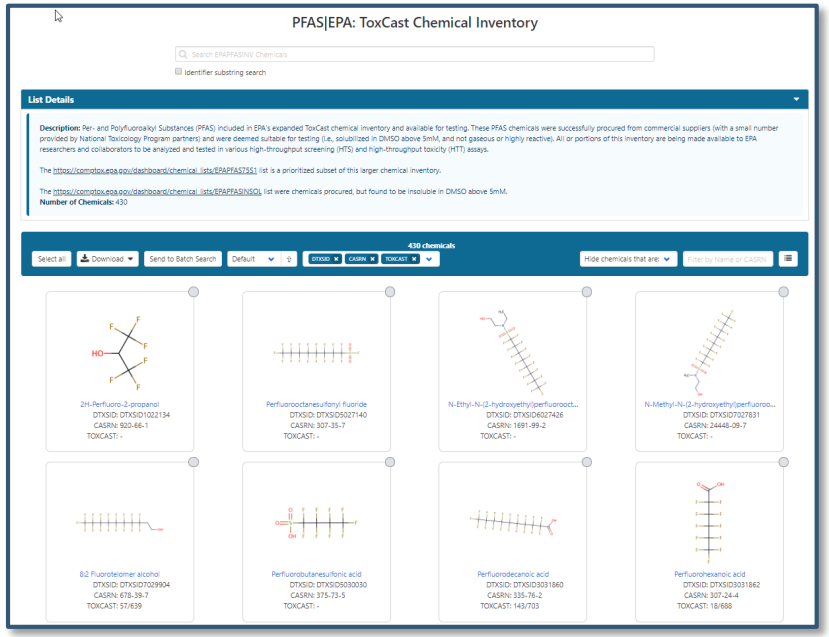
Research Area 1: *What are the human health and ecological effects of exposure to PFAS?*

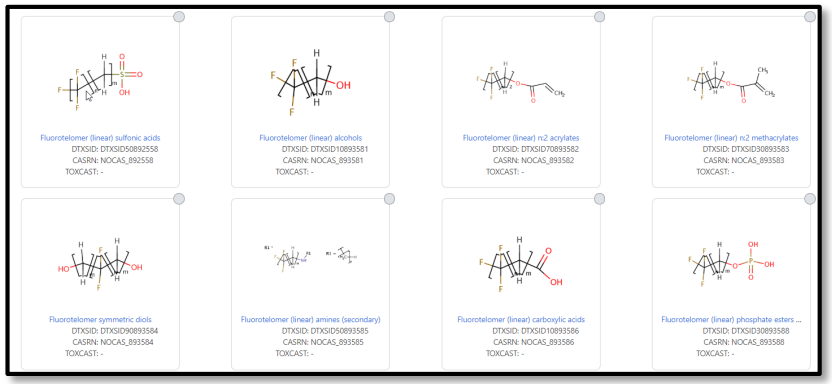
- **Using computational toxicology approaches to fill in gaps.** For the many PFAS for which published peer-reviewed data are not currently available, the EPA plans to use new approaches such as high throughput and computational approaches to explore different chemical categories of PFAS, to inform hazard effects characterization, and to promote prioritization of chemicals for further testing. These data will be useful for filling gaps in understanding the toxicity of those PFAS with little to no available data. ***In the near term,*** the EPA intends to complete assays for a representative set of 150 PFAS chemicals, load the data into the CompTox Chemicals Dashboard for access, and provide peer-reviewed guidance for stakeholders on the use and application of the information. ***In the long term,*** the EPA will continue research on methods for using these data to support risk assessments using New Approach Methods (NAMs) such as read-across and transcriptomics, and to make inferences about the toxicity of PFAS mixtures which commonly occur in real world exposures. The EPA plans to collaborate with NIEHS and universities to lead the science in this area and work with universities, industry, and other government agencies to develop the technology and chemical standards needed to conduct this research.

~1,223 PFAS currently in TSCA inventory for use in US
~ 602 of those currently active
+ unknown number of degradation and manufacturing byproducts

EPA 2019 PFAS Action Plan recognised need for approach to grouping approaches

Assembled a PFAS Chemical Library for Research and Methods Development

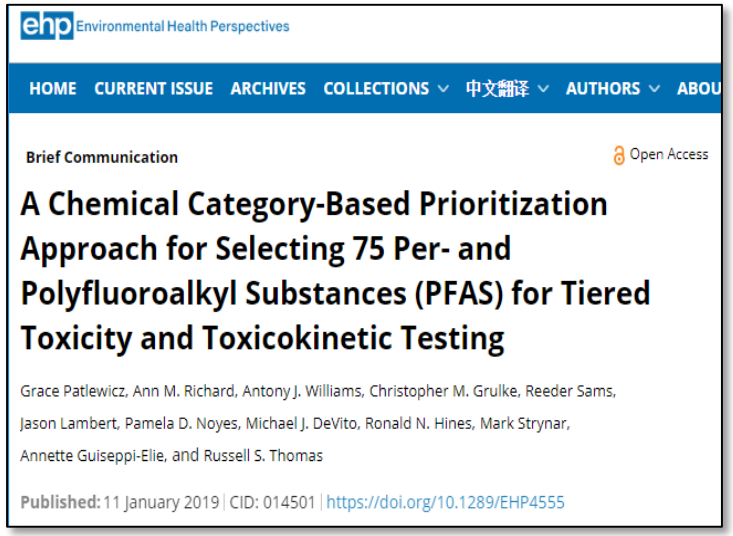




- Attempted to procure ~3,000 based on chemical diversity, Agency priorities, and other considerations
- Obtained 480 total unique chemicals
  - 430/480 soluble in DMSO (90%)
  - 54/75 soluble in water (72%) (incl. only 3 DMSO insolubles)
- Issues with sample stability and volatility
- Categories assigned based on three approaches
  - Buck et al., 2011 categories
  - Markush categories
  - OECD categories
  - Manual assignment

Kathy Coutros, Chris Grulke, and Ann Richard

Selecting a Subset of PFAS for Tiered Toxicity and Toxicokinetic Testing



ehp Environmental Health Perspectives

HOME CURRENT ISSUE ARCHIVES COLLECTIONS 中文翻译 AUTHORS ABOU

Brief Communication

Open Access

A Chemical Category-Based Prioritization Approach for Selecting 75 Per- and Polyfluoroalkyl Substances (PFAS) for Tiered Toxicity and Toxicokinetic Testing

Grace Patlewicz, Ann M. Richard, Antony J. Williams, Christopher M. Grulke, Reeder Sams, Jason Lambert, Pamela D. Noyes, Michael J. DeVito, Ronald N. Hines, Mark Strynar, Annette Guiseppi-Elie, and Russell S. Thomas

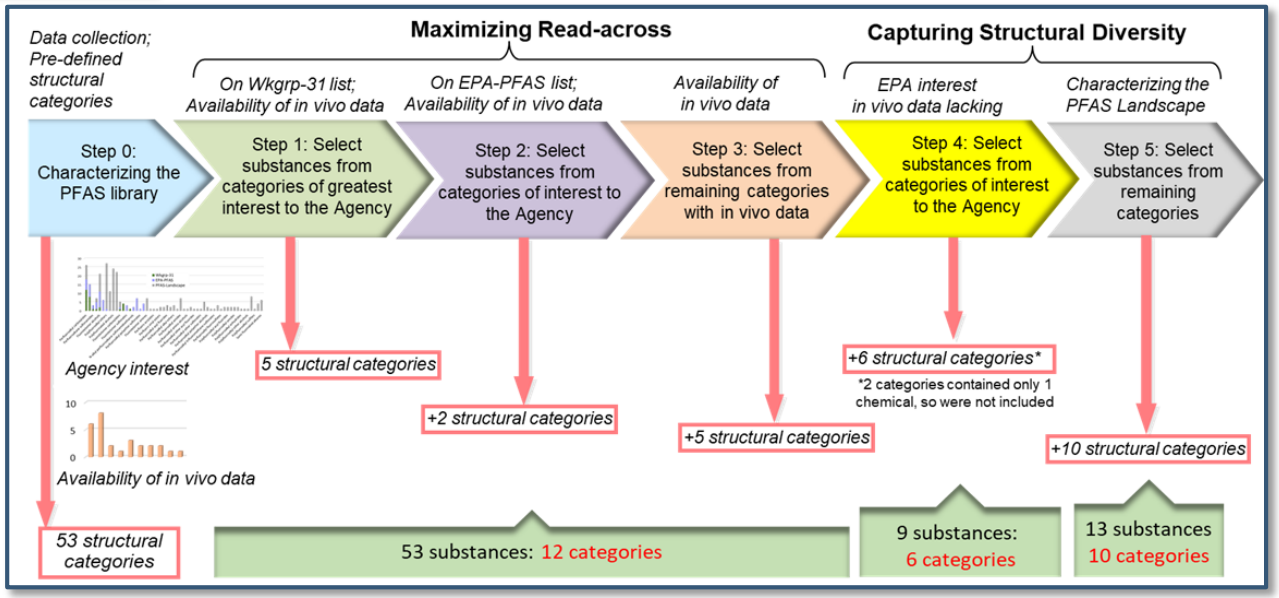
Published: 11 January 2019 | CID: 014501 | https://doi.org/10.1289/EHP4555

Goals:

- Generate data to support development and refinement of categories and read-across evaluation
- Incorporate substances of interest to Agency
- Characterise mechanistic and toxicokinetic properties of the broader PFAS landscape

Selected 150 PFAS in two phases representing 83 different categories

- 9 categories with > 3 members
- Lots of singletons



In Vitro Toxicity and Toxicokinetic Testing

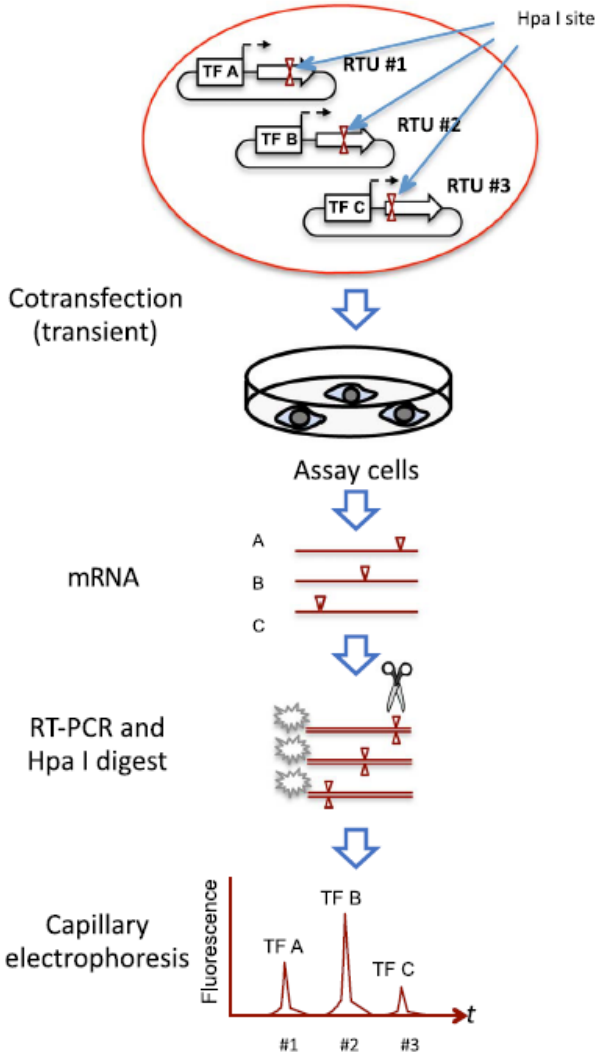
Toxicological Response	Assay	Assay Endpoints	Purpose
Hepatotoxicity	3D HepaRG assay	Cell death and transcriptomics	Measure cell death and changes in important biological pathways
Developmental Toxicity	Zebrafish embryo assay	Fertilization, lethality, and structural defects	Assess potential teratogenicity
Immunotoxicity	Bioseek Diversity Plus	Protein biomarkers across multiple primary cell types	Measure potential disease and immune responses
Mitochondrial Toxicity	Mitochondrial membrane potential and respiration (HepaRG)	Mitochondrial membrane potential and oxygen consumption	Measure mitochondrial health and function
Developmental Neurotoxicity	Microelectrode array assay (rat primary neurons)	Neuronal electrical activity	Impacts on neuron function
Endocrine Disruption	ACEA real-time cell proliferation assay (T47D)	Cell proliferation	Measure ER activity
General Toxicity	Attagene cis- and trans-Factorial assay (HepG2)	Nuclear receptor and transcription factor activation	Activation of key receptors and transcription factors involved in hepatotoxicity
	High-throughput transcriptomic assay (multiple cell types)	Cellular mRNA	Measures changes in important biological pathways
	High-throughput phenotypic profiling (multiple cell types)	Nuclear, endoplasmic reticulum, nucleoli, golgi, plasma membrane, cytoskeleton, and mitochondria morphology	Changes in cellular organelles and general morphology
Toxicokinetic Parameter	**Assay**	**Assay Endpoints**	**Purpose**
Intrinsic hepatic clearance	Hepatocyte stability assay (primary human hepatocytes)	Time course metabolism of parent chemical	Measure metabolic breakdown by the liver
Plasma protein binding	Ultracentrifugation assay	Fraction of chemical not bound to plasma protein	Measure amount of free chemical in the blood

*Assays being performed by NTP and EPA

Current work in progress

- How do the structural categories inform read-across? How are the categories enriched by the bioactivity (tiered toxicity and toxicokinetic) data being generated?

Attagene cis- and trans- Factorial Assay



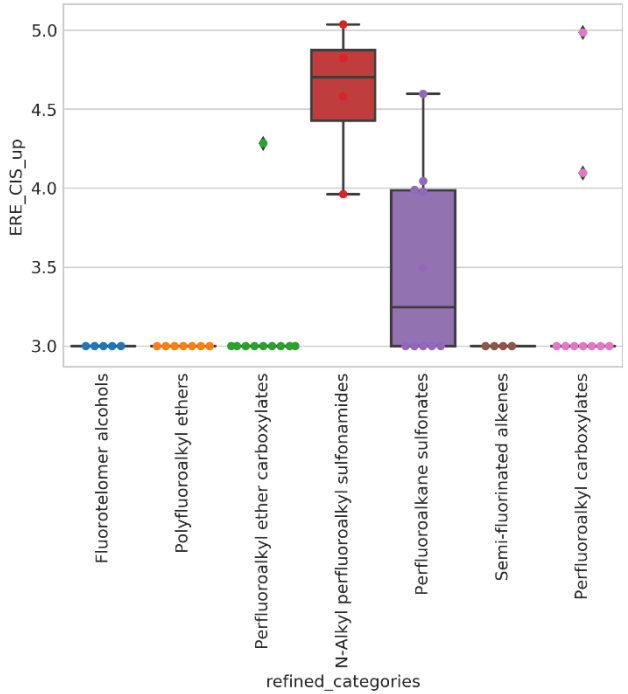
- *CIS Assay*
  - 47 Endogenous Transcription Factors
    - *Xenobiotic pathways*
    - *Cell growth/differentiation*
    - *Endocrine pathways*
    - *Stress response*
- TRANS Assay
  - 24 human nuclear receptors
  - GAL-4 formats (NR ligand-binding domains)
- HepG2 cells
  - Concentration-response testing
  - 24-hour exposure

Keith Houck and Grace Patlewicz

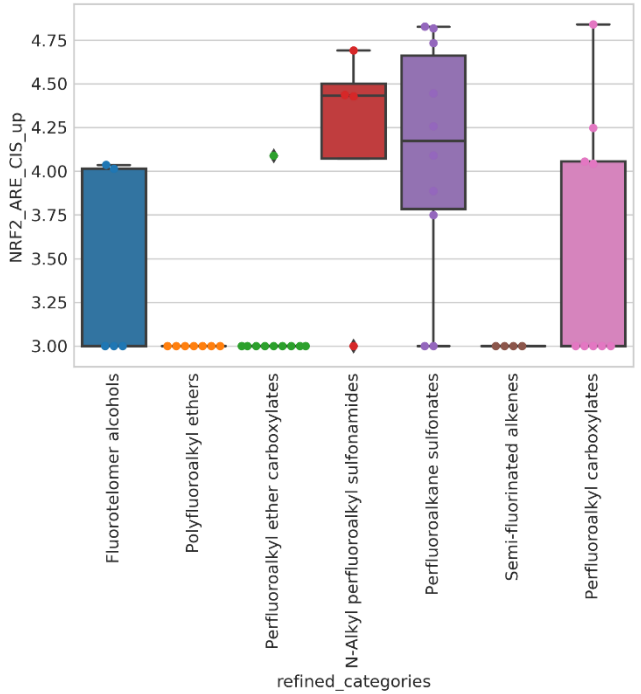
Preliminary Category-Based Analysis of the Attagene Transcription Factor Assay

United States Environmental Protection Agency

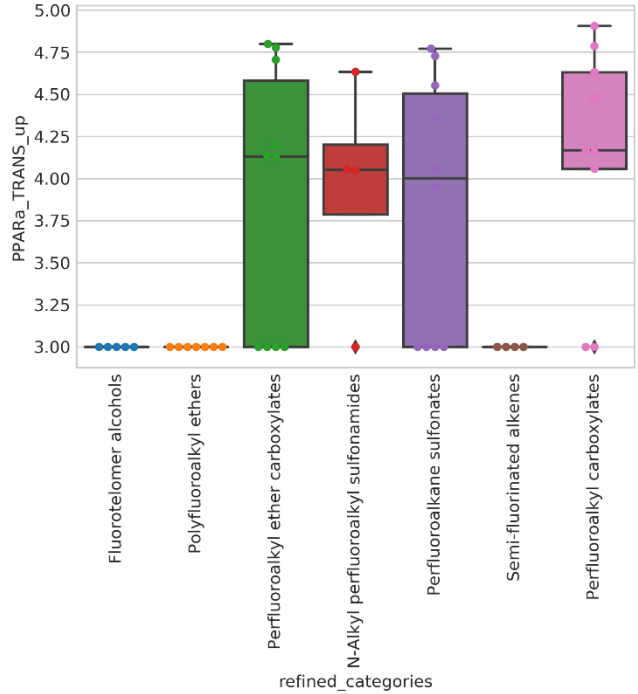
Estrogen Receptor Activity



NRF2 Activity



PPARa Receptor Activity

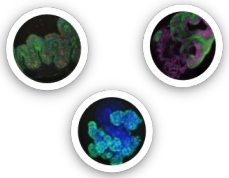


*7 categories with STD > 0.6

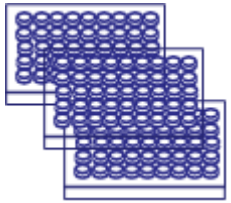
Keith Houck and Grace Patlewicz

High-Throughput Phenotypic Profiling (aka 'Cellular Pathology')

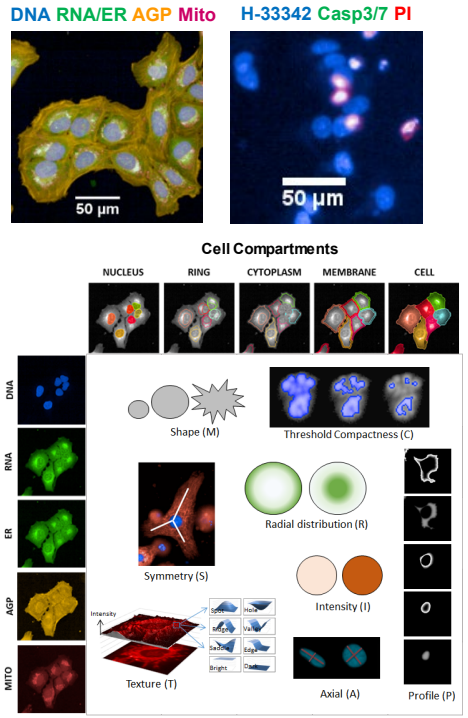
Multiple Cell Types



Concentration Response Screening

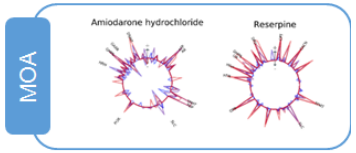


Multi-Parameter Cellular Phenotypic Profiling

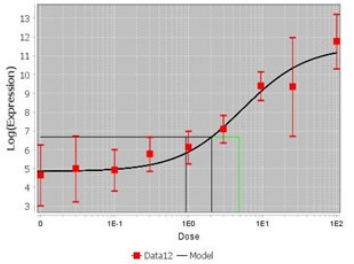


~1,300 endpoints

Mode-of-Action Inference

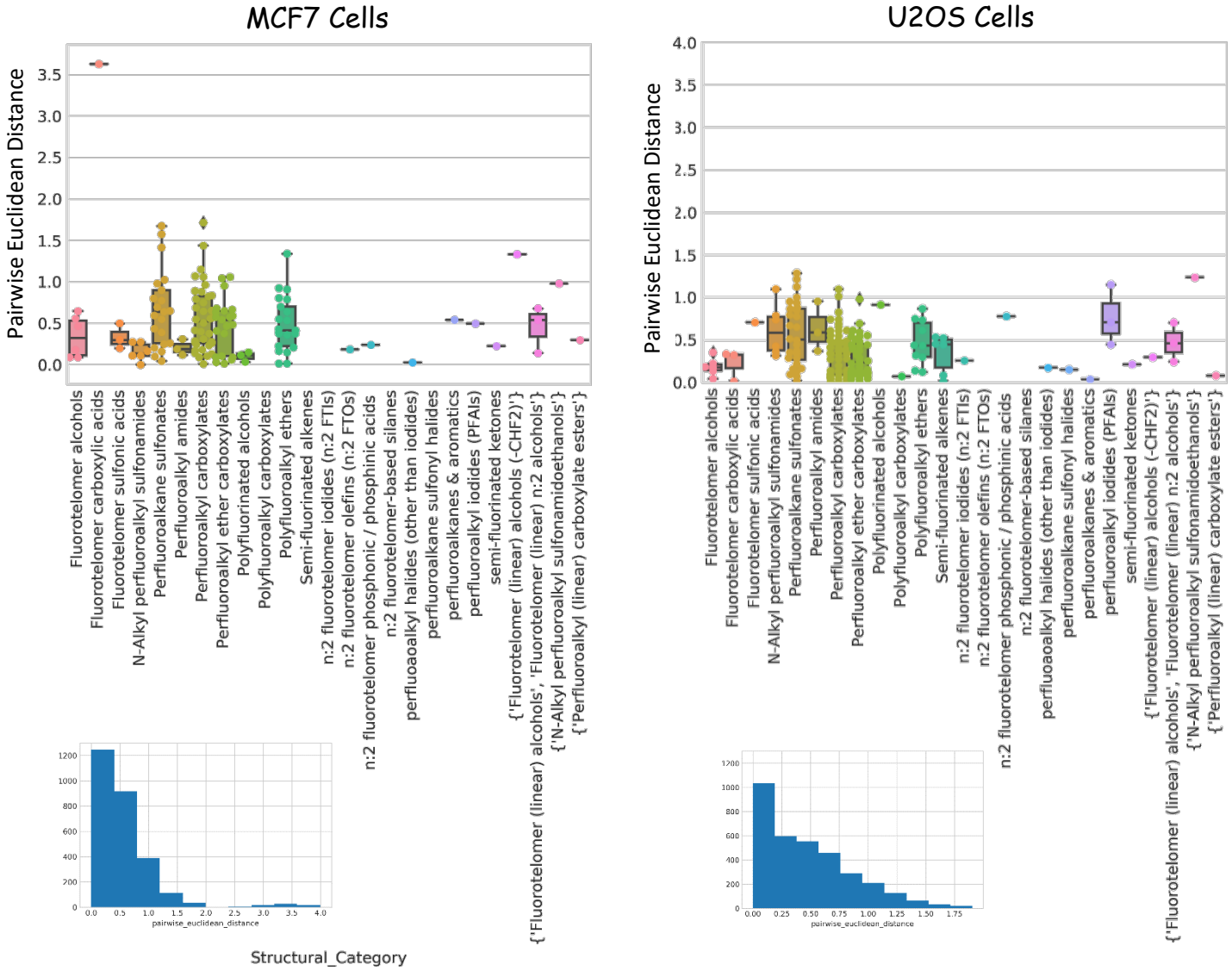


Concentration Response Modeling



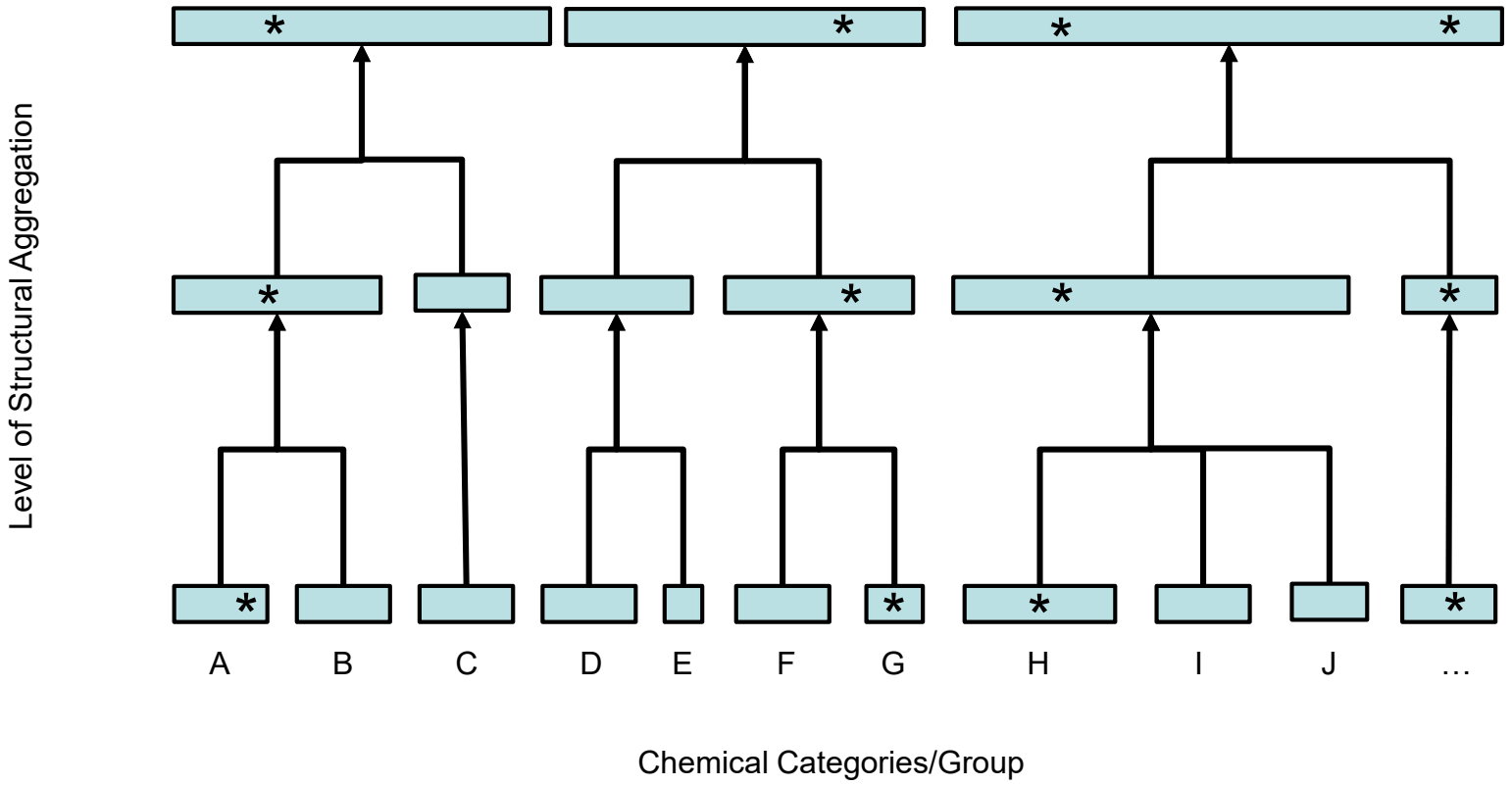
Joshua Harrill and Johanna Nyffler

Preliminary Category-Based Analysis of the Phenotypic Profiling Assay



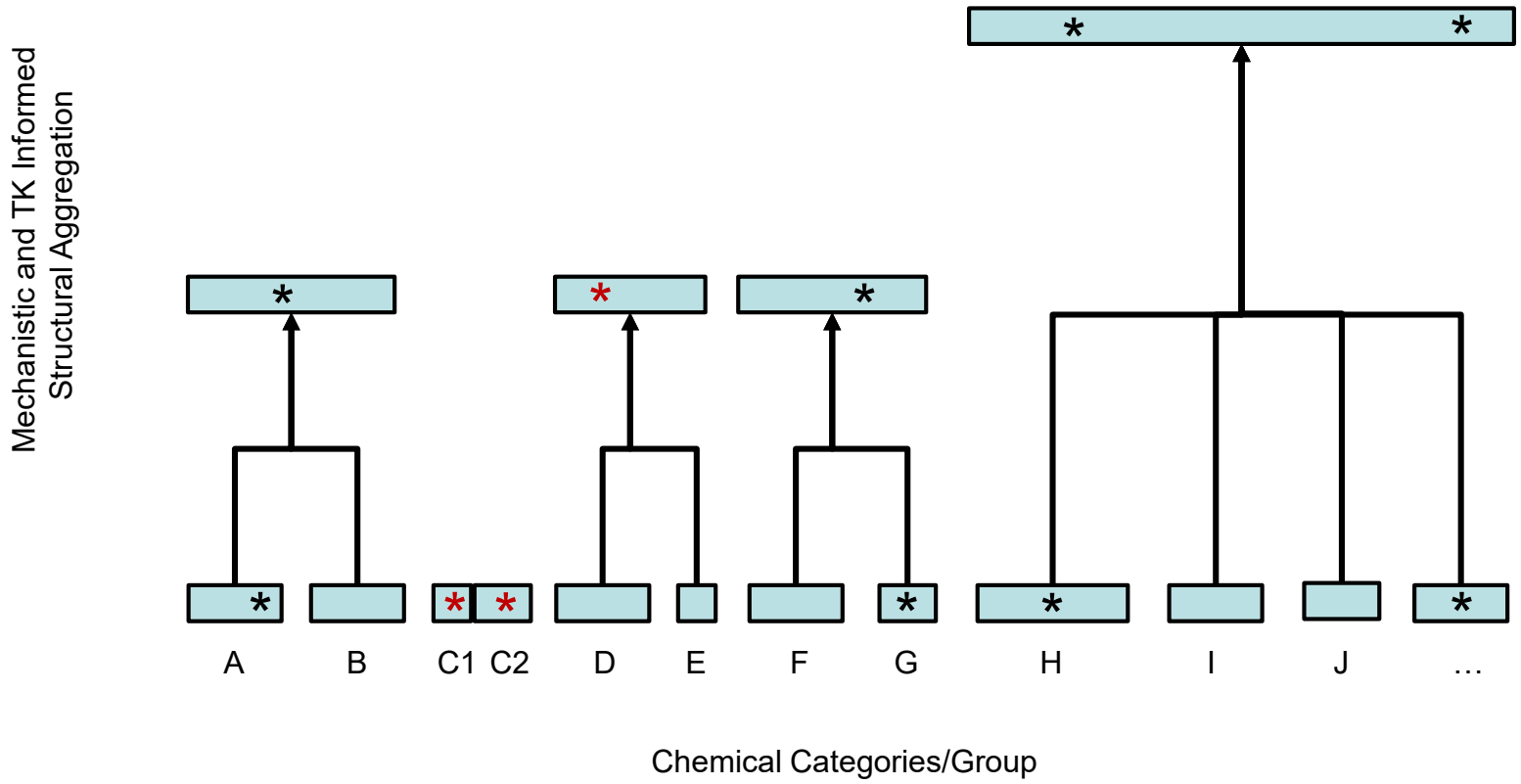
Joshua Harrill, Johanna Nyffeler, and Grace Patlewicz

Current PFAS Grouping Approaches Use Different Levels of Aggregation



* Available source *in vivo* tox study

Incorporating Mechanistic and Toxicokinetic Data to Inform PFAS Category Aggregation



*Needed *in vivo* tox study *Available source *in vivo* tox study

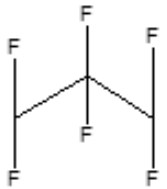
Challenges with the analysis to date…

- Initially structural category assignments were largely expert driven
- This was pragmatic based on what resources were available at the time, however it is difficult to assign membership reproducibly and objectively with a manual naming convention
- Moreover this does not facilitate profiling of other PFAS inventories/libraries of interest e.g. OECD

PFAS "Categories": Per & Poly-fluorinated alkyl substances

- "Expert"-assigned PFAS categories – manual, subjective
  - Buck et al. (DuPont), based on chemical & series informed by synthetic pathways (e.g., fluorotelomers)
  - data-gathering, occurrence reports, ecotox
  - OECD PFAS listing (>4500 chemicals) – manually assigned groupings

Poly- and Perfluorochemicals



Acyclic - Pure

Atoms: N, P, O, S, Si, Cl, Br, I = NOT

AND # of Cycles = 0

Cyclic - Pure

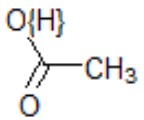
Atoms: N, P, O, S, Si, Cl, Br, I = NOT

AND # of Cycles ≥ 1

Carboxylic Acids

Atoms: N, P, S, Si, Cl, Br, I = NOT

AND



Expert category
Fluorotelomer acrylates
Fluorotelomer alcohols
Polyfluorinated alcohols
Fluorotelomer sulfonates
N-alkyl perfluoroalkyl sulfonamidoacetic acids
N-alkyl perfluoroalkyl sulfonamidoethanols
Perfluoroalkyl aldehydes
Perfluoroalkyl amides
Perfluoroalkyl carboxylates
Perfluoroalkyl acyl fluorides
Perfluoro vinyl esters
Perfluoroalkyl ketones
Semi-fluorinated alkenes
Perfluoroalkyl vinyl ethers
Perfluoroalkyl alkyl ethers
Fluorotelomer amines
Perfluoroalkyl sulfonamides
Semi-fluorinated alkanes
Perfluoroalkyl ethers
Fluorotelomer phosphates

Class	Category_Name1	Category_Name2
Alcohol	Fluorotelomer alcohols	Fluorotelomer (linear) n:2 alcohols
Sulfonic Acid	Perfluoroalkyl sulfonic acids	Perfluoroalkyl (linear C4-C10) sulfonic acids

"Expert-assigned" OECD PFAS Categories, e.g.

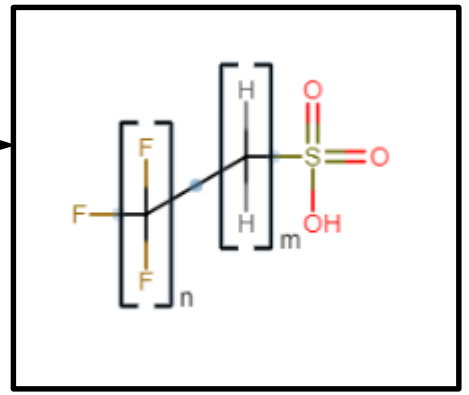
- 4730 PFAS in list
- 173 expert-assigned categories under 8 general headings (bold)
- Broad "catch-all" terms (in red)
- Structural elements, but NOT structure-based
- Requires expert to assign new chemicals to categories

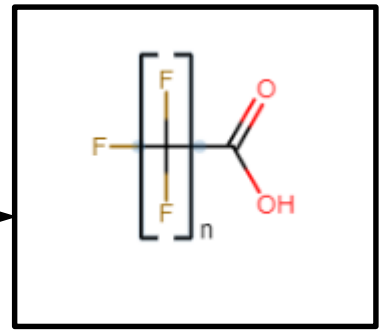
perfluoroalkyl carbonyl compounds	**CnF2n+1_C(O)_R**
perfluoroalkyl carbonyl halides	R = F/Cl/Br/I
perfluoroalkyl carboxylic acids (PFCAs), **their salts and esters**	R = OH, ONa, OCH3, etc.
other perfluoroalkyl carbonyl-based nonpolymers	to be refined
perfluoroalkyl carbonyl amides / amido ethanols **and other alcohols**	R = NH2, NH(OH), etc.
perfluoroalkyl carbonyl (meth)acrylate	R = R'_OC(O)CH=CH2
perfluoroalkyl carbonyl (meth)acrylate **polymers**	
1-H perfluoroalkyl carboxylic acids	H(CF2)nCOOH
perfluoroalkane sulfonyl compounds	**CnF2n+1_S(O)(O)_R**
perfluoroalkane sulfonyl halides	R = F/Cl/Br/I
perfluoroalkane sulfonic acids (PFSAs), their salts and esters	R = OH, ONa, OCH3, etc.
perfluoroalkane **sulfonyl-based** nonpolymers	
per- and polyfluoroalkyl ether-based compounds	**CnF2n+1_O_CmF2m+1_R**
per- **and polyfluoro**alkyl ether sulfonic acids (PFESAs), **their salts and esters, as well as derivatives**	CnF2n+1_O_CmF2m+1_SO3H
fluorotelomer-related compounds	
perfluoroalkyl iodides (PFAIs)	CnF2n+1_I
n:2 **fluorotelomer-based** non-polymers	CnF2n+1_C2H4_R, to be refined

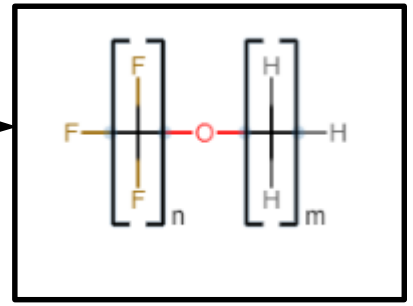
Translating Expert Categories to Markush

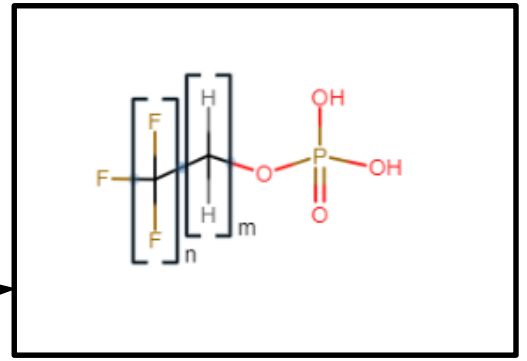
United States Environmental Protection Agency

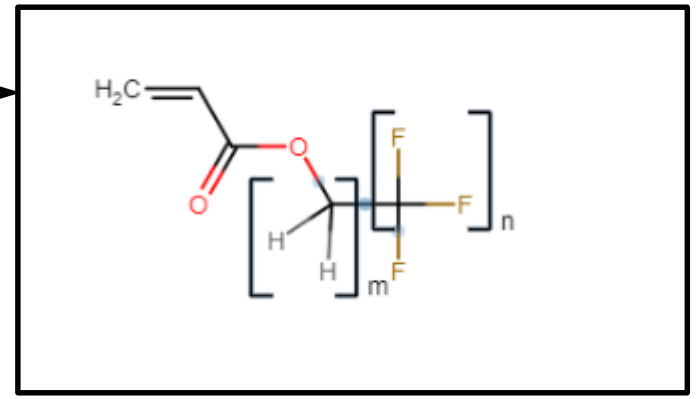
Expert category
Fluorotelomer acrylates
Fluorotelomer alcohols
Polyfluorinated alcohols
Fluorotelomer sulfonates
N-alkyl perfluoroalkyl sulfonamidoacetic acids
N-alkyl perfluoroalkyl sulfonamidoethanols
Perfluoroalkyl aldehydes
Perfluoroalkyl amides
Perfluoroalkyl carboxylates
Perfluoroalkyl acyl fluorides
Perfluoro vinyl esters
Perfluoroalkyl ketones
Semi-fluorinated alkenes
Perfluoroalkyl vinyl ethers
Perfluoroalkyl alkyl ethers
Fluorotelomer amines
Perfluoroalkyl sulfonamides
Semi-fluorinated alkanes
Perfluoroalkyl sulfonates
Perfluoroalkyl sulfonamido amines
Polyfluoroalkyl carboxylates
Perfluoroalkyl ethers
Fluorotelomer phosphates

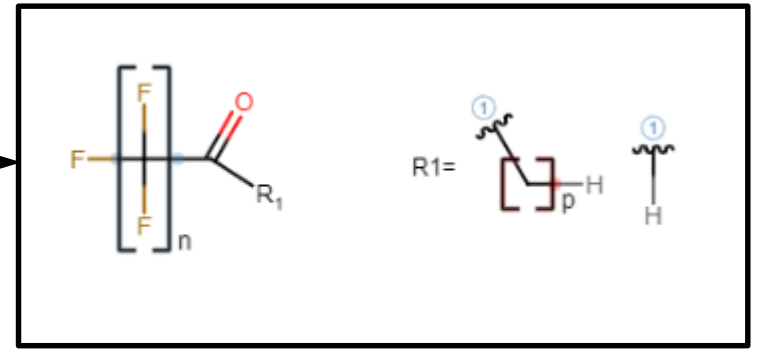


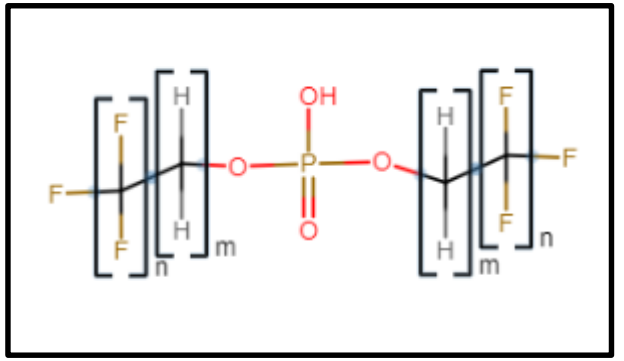












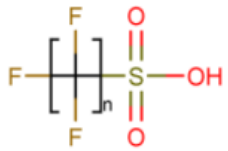
Example of Markush representation

EPA United States Environmental Protection Agency

EPA United States Environmental Protection Agency | Home | Advanced Search | Batch Search | Lists | Predictions | Downloads | Search All Data

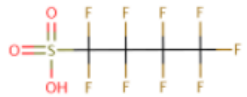
Chemistry Dashboard

Searched Chemical



Perfluoroalkyl sulfonates
DTXSID: DTXSID70892979
CASRN: NOCAS_892979

markush



Perfluorobutanesulfonic acid
DTXSID: DTXSID5030030
CASRN: 375-73-5

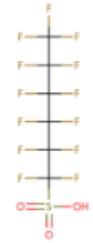
markush

Perfluorooctanesulfonic acid
DTXSID: DTXSID3031864
CASRN: 1763-23-1

markush

Perfluorodecanesulfonic acid
DTXSID: DTXSID3040148
CASRN: 335-77-3

markush

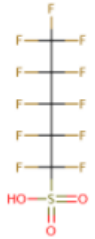


Perfluorohexanesulfonic acid
DTXSID: DTXSID7040150
CASRN: 355-46-4

markush

Perfluoroheptanesulfonic acid
DTXSID: DTXSID8059920
CASRN: 375-92-8

markush



Perfluoropentanesulfonic acid
DTXSID: DTXSID8062600
CASRN: 2706-91-4

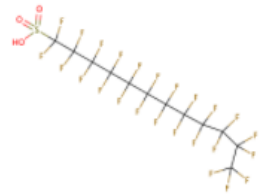
markush

Perfluorononanesulfonic acid
DTXSID: DTXSID8071356
CASRN: 68259-12-1

markush

Perfluoropropanesulfonic acid
DTXSID: DTXSID30870531
CASRN: 423-41-6

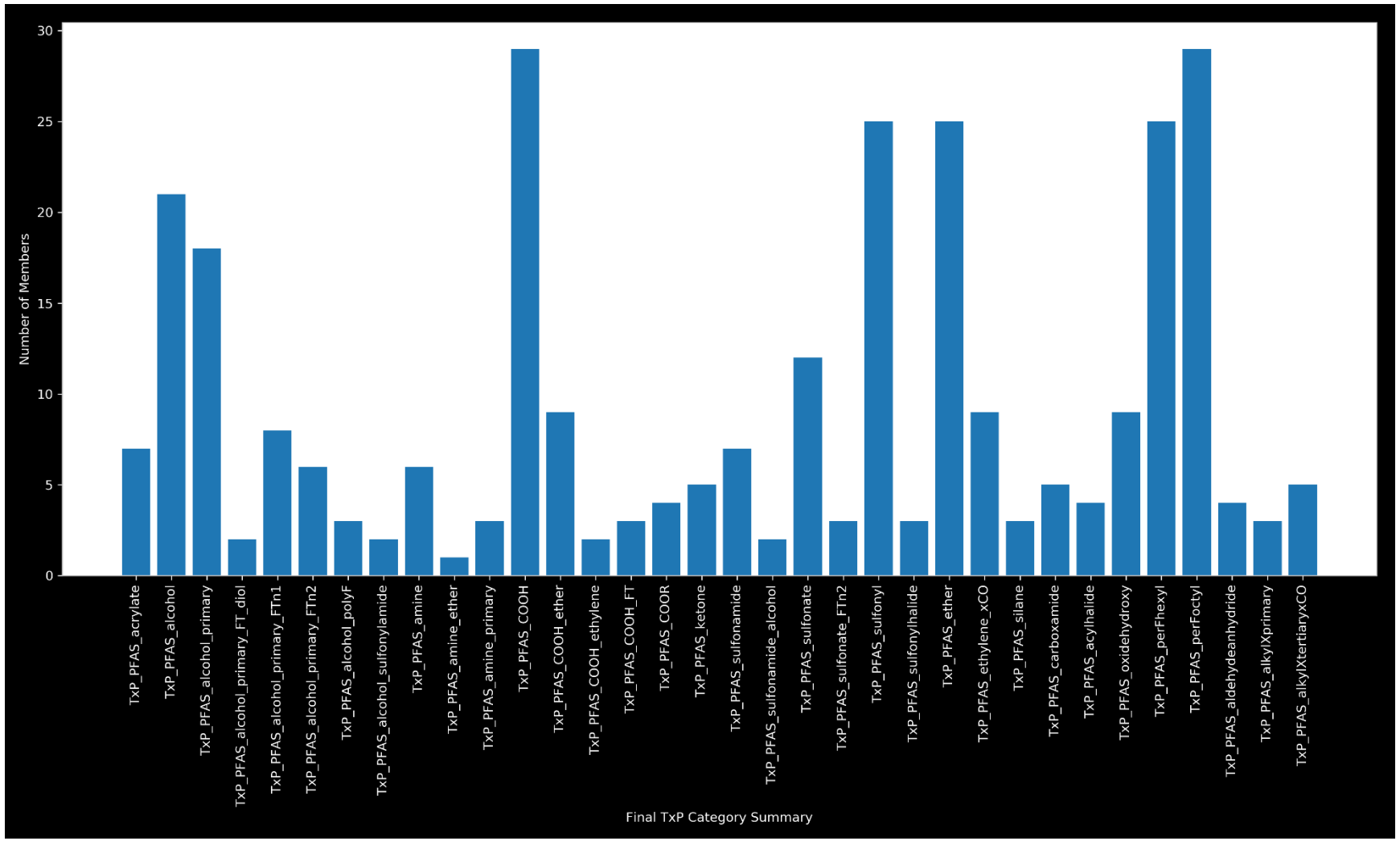
markush



Perfluorododecanesulfonic acid (PFDOS)
DTXSID: DTXSID20873011
CASRN: 79780-39-5

Exploiting fixed fingerprints to facilitate objective structural categories

- For the ~150 set, have aimed to harmonise the 3 schemes using fixed ToxPrints
- Defined rules on membership based on specific features
- Extendable to incorporate other information i.e. bioactivity



Ann Richard and Grace Patlewicz

Take home messages

- Computational toxicology approaches impact many aspects of regulatory contexts
- Outlined how computational approaches fit within an IATA
- Illustrated how we have explored coupling TTC & HTE for a risk-based prioritisation application
- Discussed read-across approaches, tools & their frameworks
- Proposed a harmonised framework for read-across approaches

Take home messages

- Outlined GenRA, how it was developed and how it is aligned with this framework – public tool
- Initial GenRA (baseline) considers structural similarity but current work has evaluated the quantitative impact of physicochemical similarity (as it relates to bioavailability) and transitioning to dose predictions e.g. acute toxicity LD50
- Highlighted the research efforts of using chemical structural groupings to underpin selection of representative PFAS for toxicity and toxicokinetic testing using NAMs

Acknowledgements

- Many but in particular..
- Imran Shah
- George Helman
- Tony Williams
- Richard Judson
- Ann Richard
- Chris Grulke
- Keith Houck
- Jason Lambert
- John Wambaugh
- Joshua Harrill
- Johanna Nyffeler
- Rusty Thomas