Transitioning towards objective Read-across approaches: landscape, research, and practical application

Grace Patlewicz, PhD

National Center for Computational Toxicology (NCCT)

US Environmental Protection Agency (US EPA)

RTP, NC

patlewicz.grace@epa.gov

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

Conflict of Interest Statement: NONE.

Abbreviations/Definitions

- **Target – substance of interest, data poor**
- **Source – analogue with data which will be used to make the read-across prediction**
- **PMN – Premanufacture notice**
- **PPRTV - Provisional Peer Reviewed Toxicity Values (for Superfund)**
- **GenRA – Generalised Read-across**

Talk Objectives

Understanding:

- **Definitions of read-across, category & analogue approaches**
- **Read-across development and assessment frameworks**
- **Harmonised framework for read-across**
- **Selected read-across tools**
- **Ongoing issues with read-across**
- **Current directions towards quantifying read-across performance and its associated uncertainties given ‘big data’ needs**
- **Generalised Read-across (GenRA) – an approach and an application**

Talk Outline

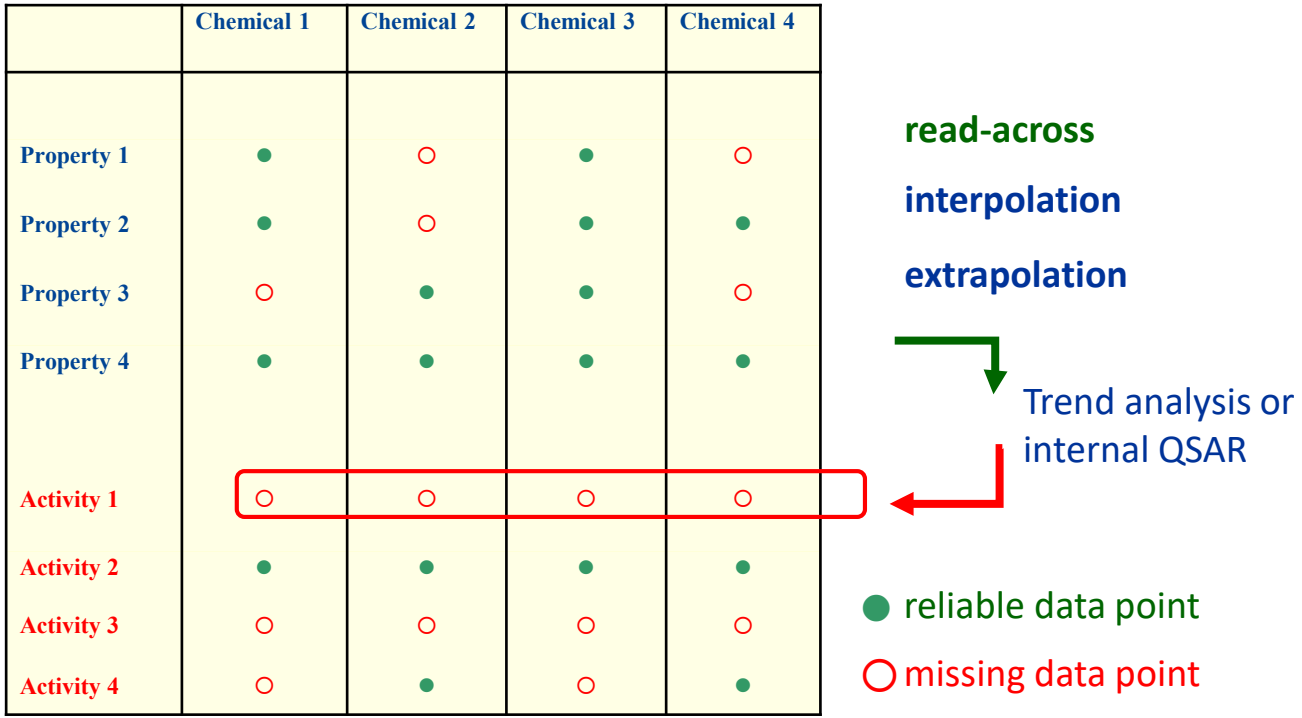
- Definitions
- Frameworks for read-across development and assessment
- Harmonised hybrid read-across framework
- Selected tools for read-across
- Ongoing issues with read-across and its acceptance
- Current directions towards quantifying read-across performance and its associated uncertainties given ‘big data’ needs
- Generalised Read-across (GenRA) – an approach and an application

Definitions: Chemical grouping approaches

- Read-across describes one of the techniques for filling data gaps in either the analogue or category 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).

Uses of Read-across



Uses of Read-across

- Examples where “read-across” approaches are applied include:
  - US EPA Provisional Peer Reviewed Toxicity Values (PPRTVs) where data is lacking for a specific substance of interest
  - EPA Test Rules – Industry registrants providing information to satisfy a test rule
  - EPA Pre Manufacture Notifications (PMN) – QSARs such as those in Epiwin and ECOSAR are routinely used for e-fate and ecotox predictions but read-across is relied upon for non cancer endpoints
  - ASTDR Emergency response values – an accidental spill that requires an immediate assessment of acute toxicity for first responders
  - REACH registrations – addressing information requirements

Developing a read-across assessment

- Existing guidance and resources that can be helpful in developing a read-across assessment:
  - Technical regulatory guidance has been published by OECD and ECHA
  - OECD guidance from 2007 was updated in 2014
  - ECHA Chapter 6 QSARs and Grouping of Chemicals as well as practical guides
- However, many papers have been published that complement and augment the regulatory guidance for development of read-across
  - Wang et al (2012) Application of computational toxicological approaches in human health risk assessment. I A tiered surrogate approach (EPA PPRTVs)

Developing a read-across assessment

- Selected literature include:
  - ECETOC TR116 category approaches, Read-across, (Q)SAR
  - Wu et al (2010) – Framework for using structural, reactivity, metabolic and physicochemical similarity to evaluate suitability of analogs for SAR based toxicological assessments
  - Patlewicz et al (2013) Use of category approaches, read-across and (Q)SAR general considerations
  - Patlewicz et al (2015) Building scientific confidence in the development and evaluation of read-across
  - Ball et al (2016) Towards Good Read-across Practice

Summary highlights of read-across development frameworks

Framework	ECHA	OECD	Wu et al	Wang	Patlewicz
Context	REACH	International regulatory purposes	Product Stewardship	Quantitative risk assessment	Regulatory purposes/Product stewardship
Approach	Analogue/Category – aim is to fill an endpoint specific study. Focused on structural similarity as a starting point Approach is more hypothesis driven	Analogue/Category – a generalisation of the ECHA approach	Analogue Systematic stepwise evaluation of analogue suitability based on structure, reactivity, p-chem and metabolism	Analogue Approach is based on a WOE assessment from structure, ADME and toxicity considerations	Analogue Stepwise approach considering general (pchem, reactivity, metabolism) and endpoint specific considerations
Terms of reference	Target/Source	Target/Source	Substance of interest/Analogue	Chemical of Concern/Surrogate	Analogue/Category
Scope	Endpoint specific	Endpoint specific	Systematic stepwise evaluation of analogue suitability based on structure, reactivity, p-chem and metabolism Most sensitive/relevant endpoint – focused on repeated dose toxicity endpoints; quantitative risk assessment	Approach is based on a WOE assessment from structure, ADME and toxicity considerations. "Best" surrogate is selected from a set of candidates based on most similar and most conservative toxicity value	Approach is aimed to identify source analogues that can be used to address as many endpoints as appropriate, even though the read-across prediction itself is justified on an endpoint per endpoint basis and some source analogues might be excluded from the prediction itself if they are not appropriate for specific endpoints of

Reviewed in Patlewicz et al., 2018

Ongoing issues with read-across

- Although there is much guidance for developing read-across assessment, acceptance still remains an issue, especially for regulatory purposes.
- A key issue thwarting acceptance relates to the “uncertainty of the read-across”
- As such there have been many efforts to identify the sources of uncertainty in read-across, characterise them in a consistent manner and identify practical strategies to address and reduce those uncertainties.
- Notable in these efforts have been the development of frameworks for the assessment of read-across. These allow for a structured assessment of the read-across justification.

Sources of uncertainty in read-across

- Analogue or category approach? (no. of analogues)
- Completeness of the data matrix – no. of data gaps
- Data quality for the underlying analogues for the target and source analogues
- Consistency of data across the data matrix – concordance of effects and potency across analogues
- Overarching hypothesis/similarity rationale – how to identify similar analogues and justify their similarity for the endpoint of interest
- Address the dissimilarities and whether these are significant from a toxicological standpoint e.g. ToxDelta
- Presence vs. absence of toxicity
- Toxicokinetics

Frameworks for Assessing Read-across

- Blackburn & Stuard
- Patlewicz et al (2015)
- Schultz et al (2015)
- ECHA RAAF (2015, 2017)

- These aim to identify, document and address the uncertainties associated with read-across inferences/predictions

Frameworks for the assessment of read-across

- Schultz et al (2015)
- Outlined a strategy for structuring and reporting a read-across
- Defined different read-across scenarios
- Two main aspects tackled:
  - an assessment of the similarity of the source analogues
  - an assessment of the mechanistic relevance and completeness of the read-across (number of analogues, absence/presence of toxicity, quality of underlying data, temporal and dose response relationship between mechanistically relevant endpoints
- Three scale grading of the overall read-across confidence Low, Medium, High

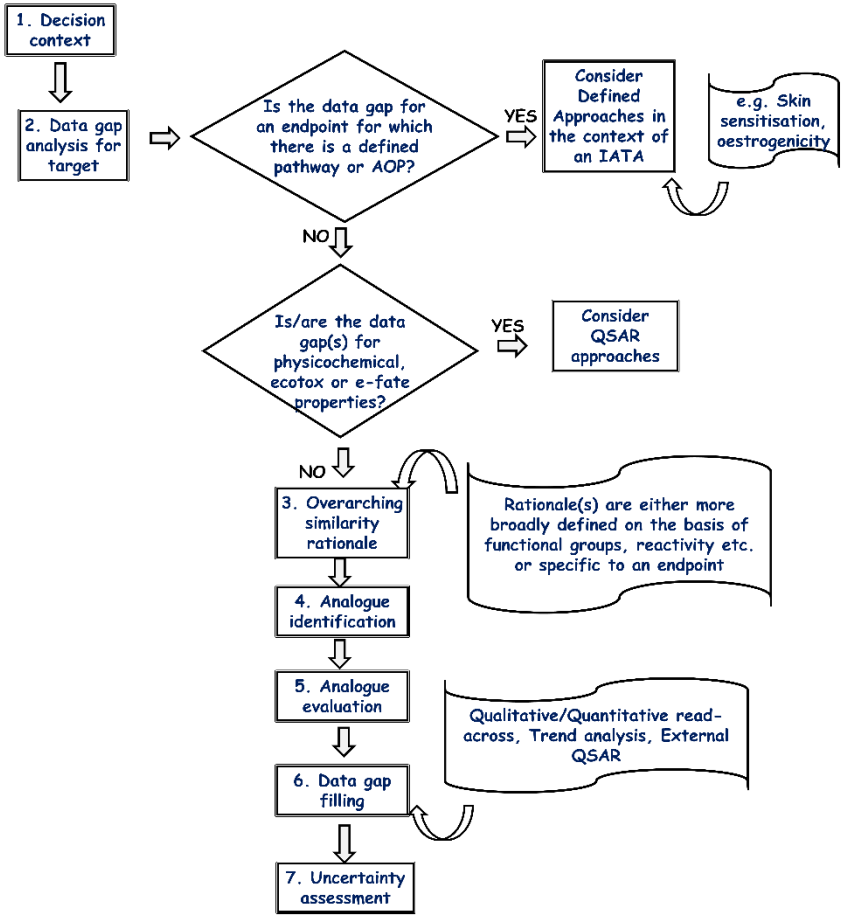
Frameworks for the assessment of read-across: RAAF

- Six scenarios identified
- For each scenario there will be a number of scientific considerations
- Each is associated with an “assessment element” (AE)
- Each AE is scored from 1-5 where 5 is “acceptable with high confidence” to 1 is not acceptable
- These scores are termed Assessment Options (AO)
- A minimum score of 3 is needed for a read-across to be taken up and used to inform decision making
- There are common assessment elements e.g. reliability of the underlying data and there are scenario specific elements e.g. common underlying mechanism for scenario 2

Summary highlights of read-across assessment frameworks

Framework	ECHA	Blackburn and Stuard	Patlewicz et al	Schultz et al
Context	REACH	Product Stewardship	Regulatory purposes & Product stewardship	Regulatory purposes & Product stewardship
Scope	Analogue/Category	Analogue/Category	Analogue/Category	Analogue/Category
Framework	Scenarios addressing analogue (2) and category (4) approaches as described above Each scenario is associated with a number of assessment elements (AE) (both common and scenario specific).	Framework addresses 3 aspects: analogue suitability (covered in Wu et al, 2010); data quality of the analogues; consistency of the data across the analogues and relative to the target	Identifies the sources of uncertainty in relationship to the data and similarity context	Different scenarios are articulated to frame up to 11 different similarity criteria, factors proposed to evaluate mechanistic relevance and completeness of the read-across

A harmonised hybrid read-across workflow



Patlewicz et al., 2018

Ongoing issues with read-across

- These frameworks allow for a structured assessment of the read-across justification.
- The next step is how those uncertainties can be addressed
- Blackburn and Stuard (2014) propose the use of assessment factors
- The RAAF and the work by Schultz et al (2015) advocate the use of New Approach Methods (NAM) (e.g. High Throughput Screening (HTS) data) to enhance the scientific confidence of a read-across
- Examples have been published by Schultz (2017) and colleagues
- Others such as Shah et al (2016) or Zhu et al (2016) have explored quantifying the uncertainties of read-across and using NAM data (e.g. big data) in conjunction with chemical structure information in a 'QSAR-like' read-across (Generalised Read-Across (GenRA)
- Some of these efforts have been implemented into read-across tools

Selected read-across tools

Tool	AIM	ToxMatch	AMBIT	OECD Toolbox	CBR A	ToxRead	GenRA
Analogue identification	X	X	X	X	X	X	X
Analogue Evaluation	NA	X	X by other tools available	X	X	X For Ames & BCF	NA
Data gap analysis	NA	X	X Data matrix can be exported	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

Quantifying uncertainty & Assessing performance of read-across

- **GenRA (Generalised Read-Across) is a “local validity” approach**
- **Predicting toxicity as a similarity-weighted activity of nearest neighbours based on chemistry and bioactivity descriptors**
- **Systematically evaluates read-across performance and uncertainty using available data**

Jaccard similarity:

GenRA - Approach

I. Data

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

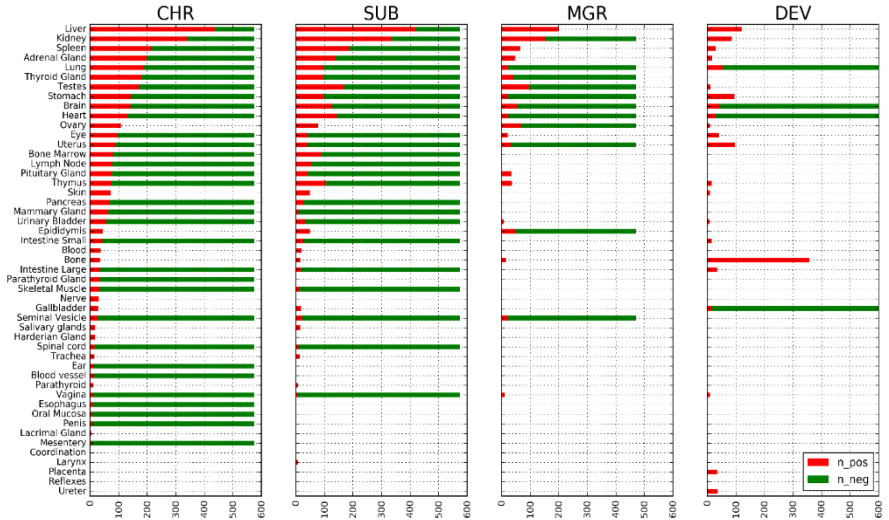
574 toxicity effects (tox) ToxRefDB

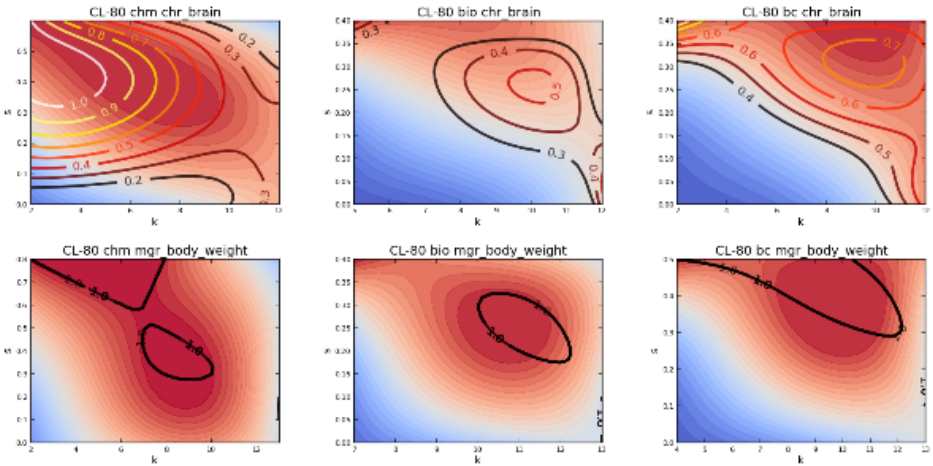
II. Define Local neighbourhoods

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

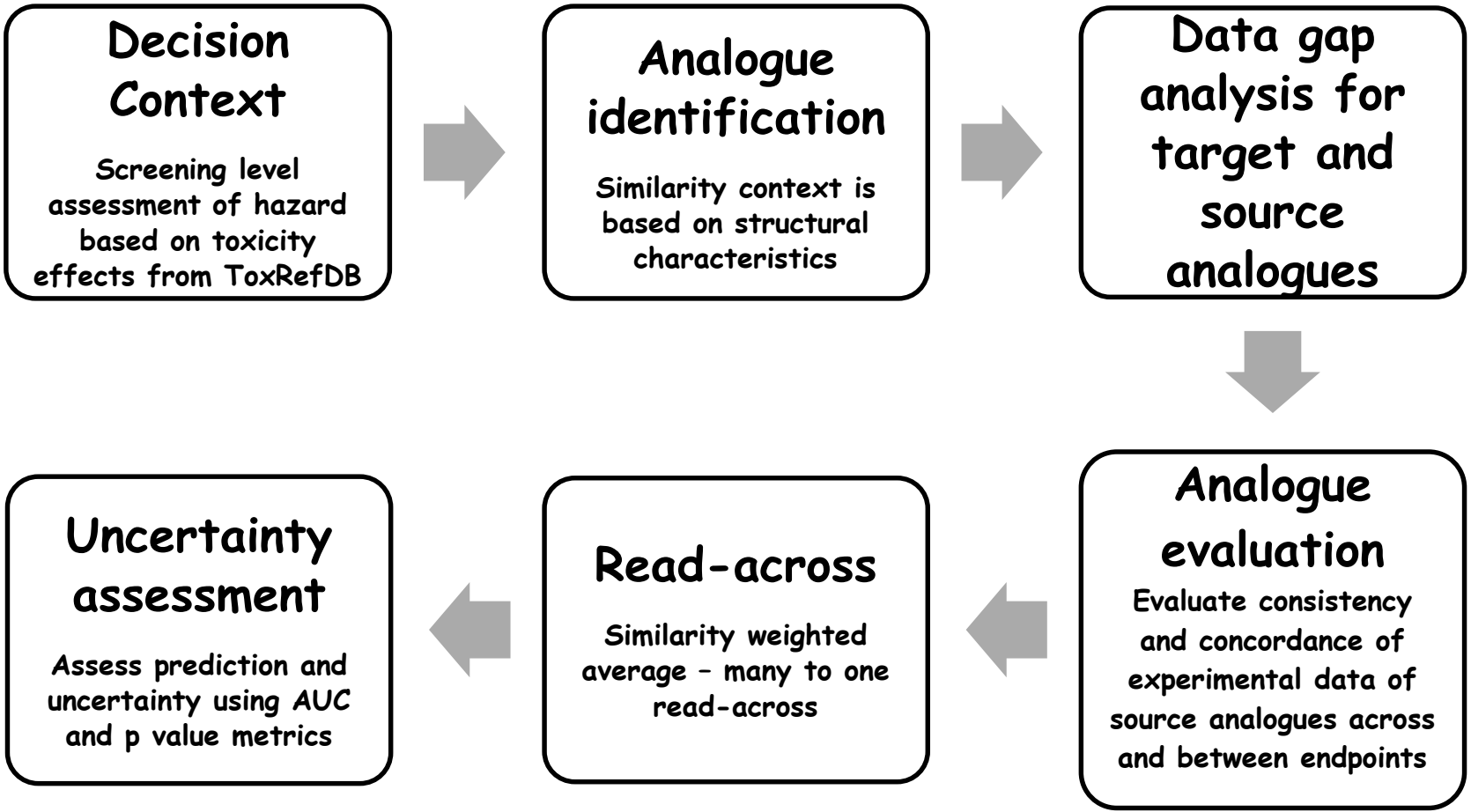
III. GenRA

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



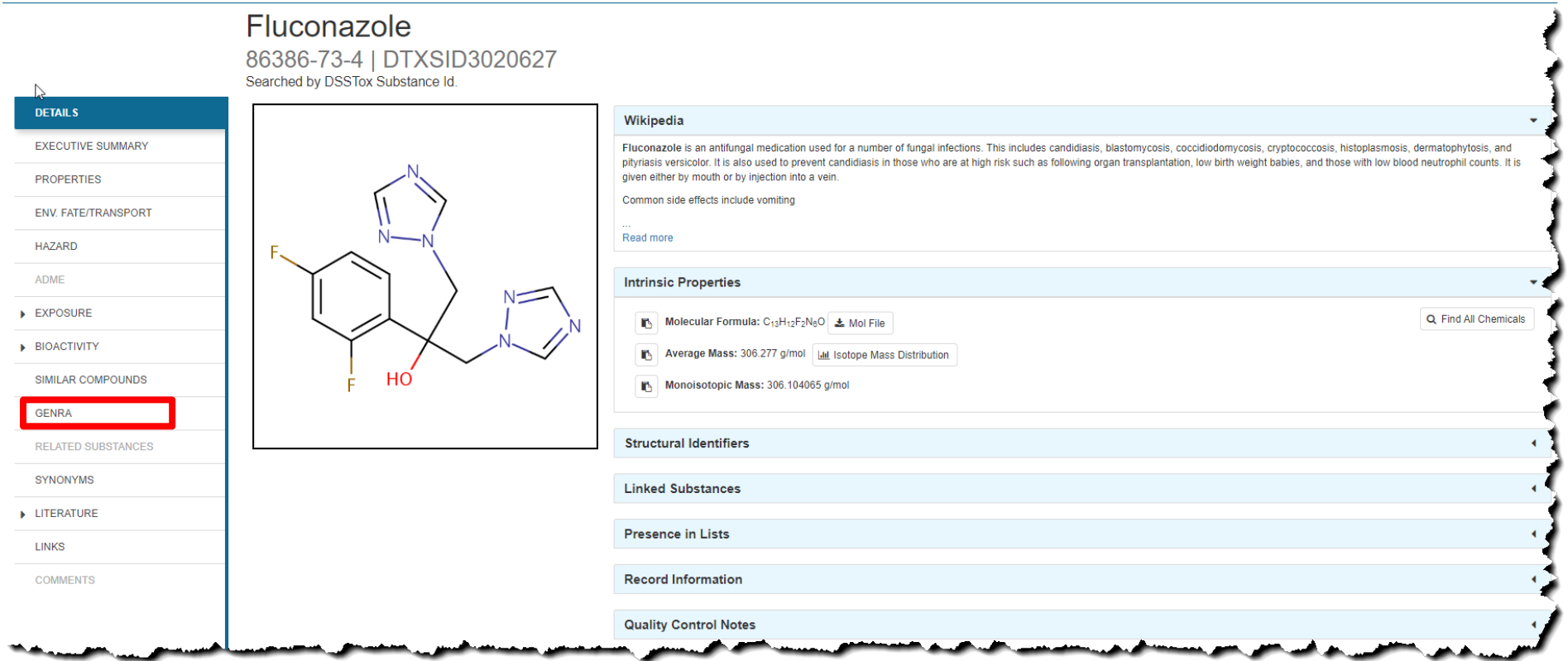


Read-across workflow in GenRA



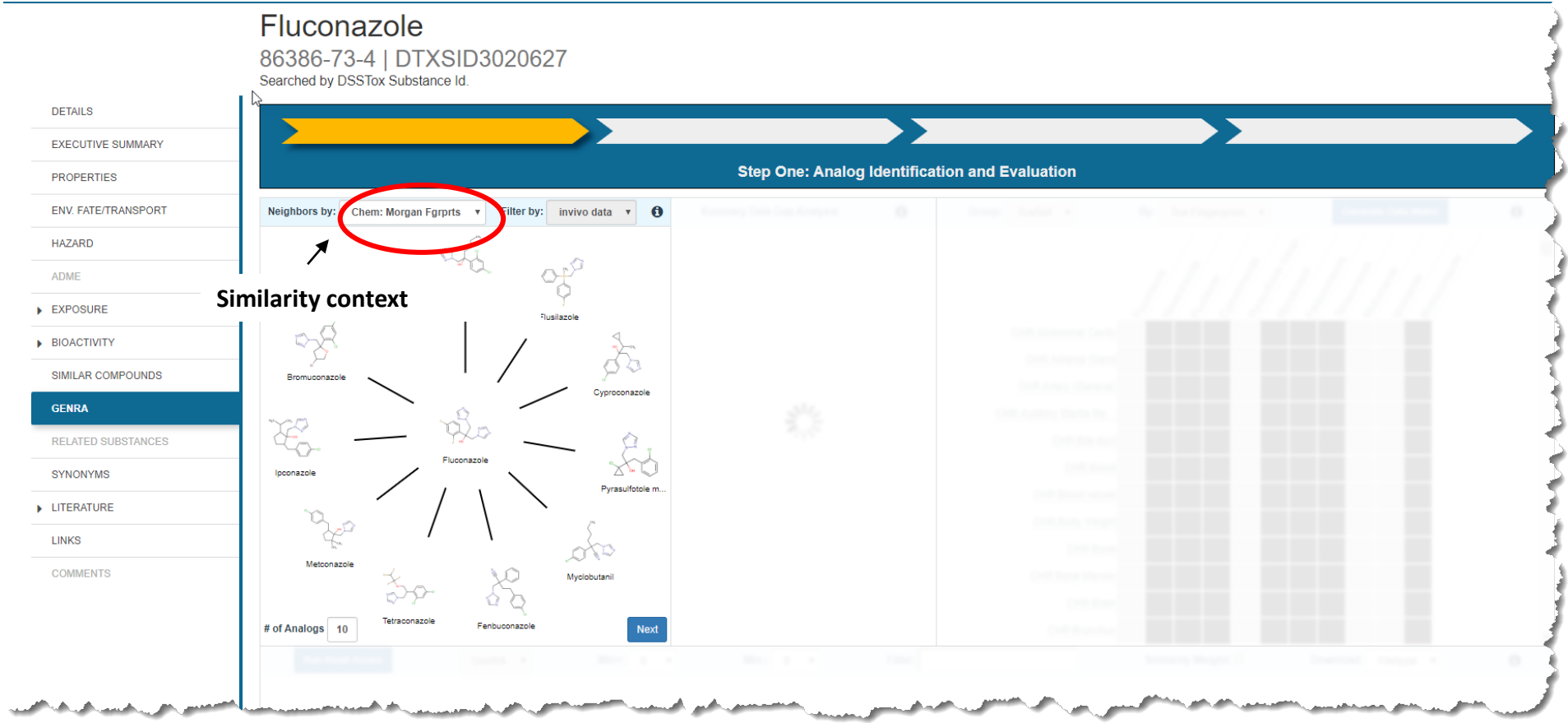
GenRA tool in reality

- **Integrated into the EPA CompTox Chemicals dashboard**

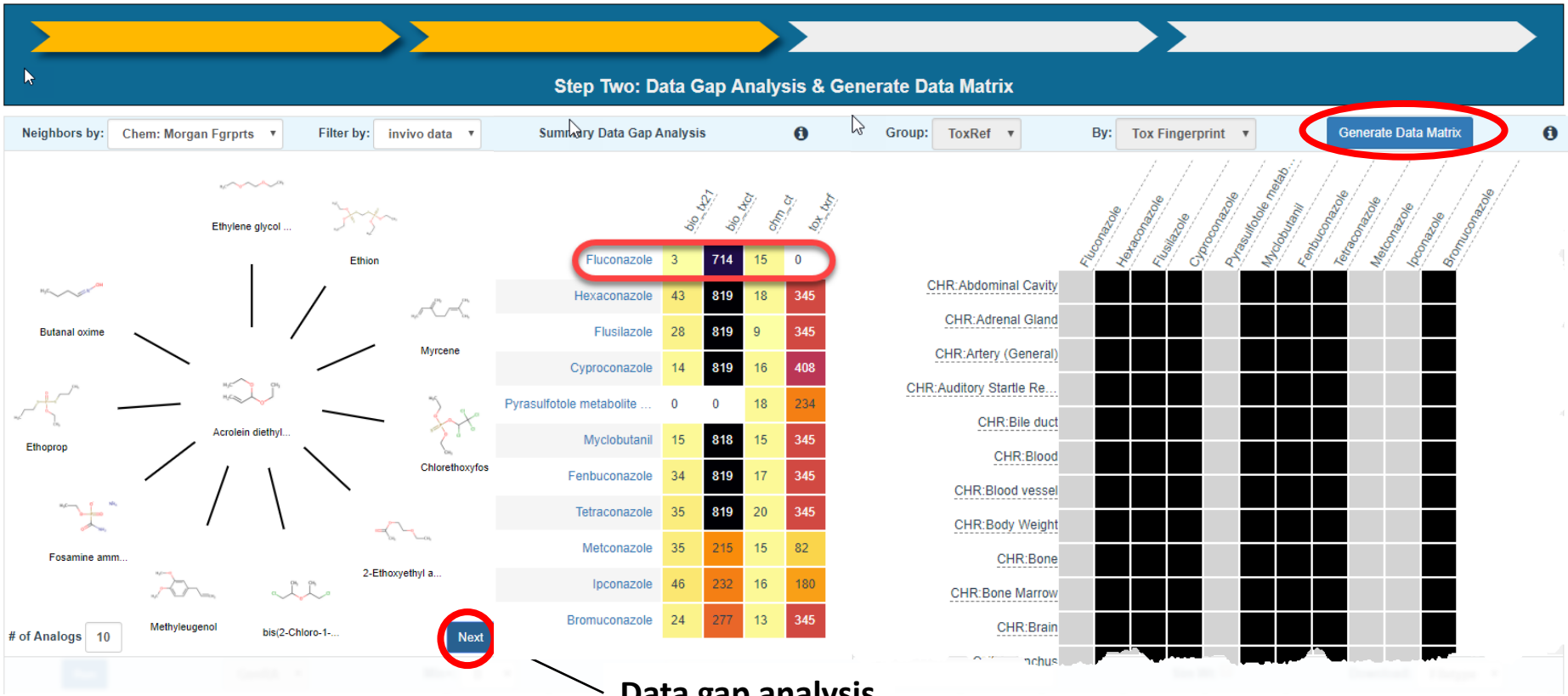


GenRA tool in reality

- **Structured as a workflow**



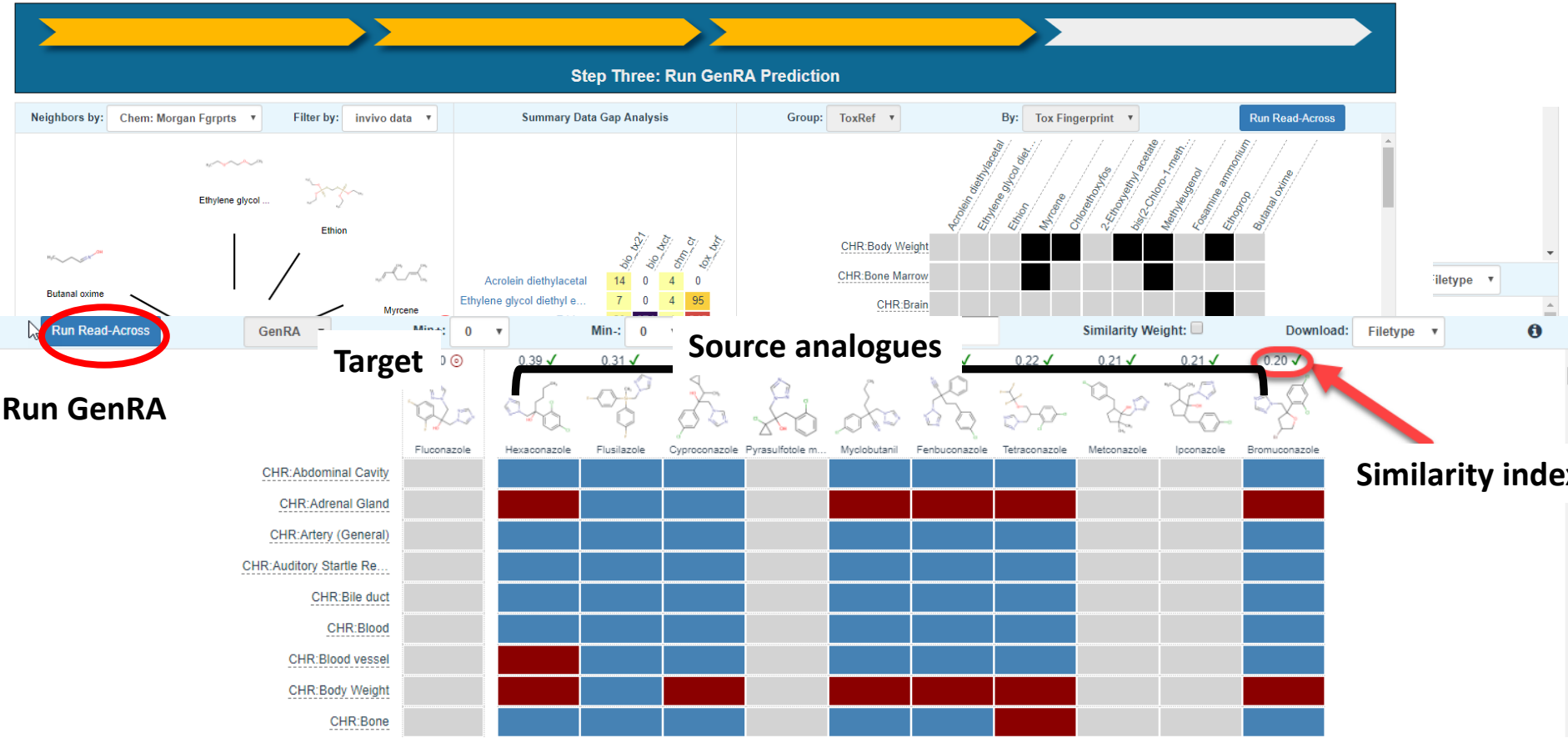
GenRA tool in reality



Data gap analysis

GenRA tool in reality

GenRA



GenRA – Next Steps

- Ongoing research:
- Summarising and aggregating the toxicity effect predictions to guide end users – what are the effects to be concerned about and which effect predictions are we most confident about
- Consideration of other information to define and refine the analogue selection – e.g. physicochemical similarity, metabolic similarity, reactivity similarity, bioactivity similarity, transcriptomics similarity…
  - Quantifying the impact of physicochemical similarity on read-across performance
  - Quantifying the impact of transcriptomic similarity on read-across performance

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 first with quantitative classical toxicity data – e.g. acute rat oral toxicity, ToxRefDB v2
  - In the future, bringing in quantitative HTTr data

GenRA & Physchem Similarity Context

- Important context of similarity in read-across
- Models “bioavailability”
- Properties selected: Lipinski Rule of 5 (LogP, MW, # HB donors/acceptors)
- Two approaches investigated as a means to identify source analogs and evaluate their predictive performance relative to GenRA:

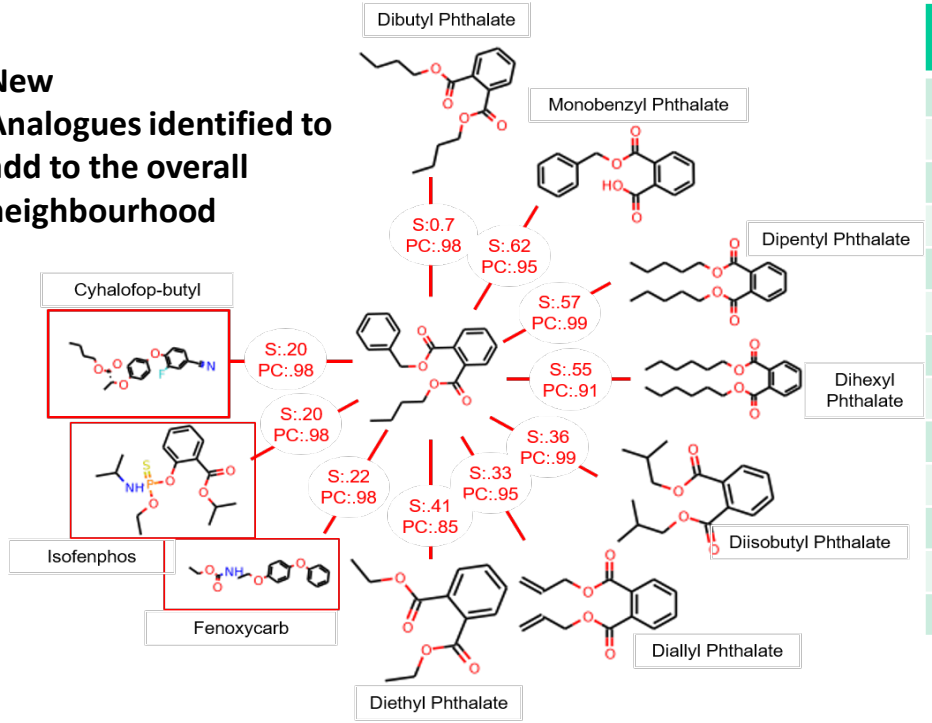
Approach 1: “Filter”	Approach 2: “Search Expansion”
Subcategorise from a set of analogues identified based on structural similarity	“Frontload” both structure and physchem into analogue identification
‘Common’ approach	‘Novel’ approach

Helman et al., 2018

Case Study: Butyl Benzyl Phthalate

Approach 2: Search Expansion

New Analogues identified to add to the overall neighbourhood

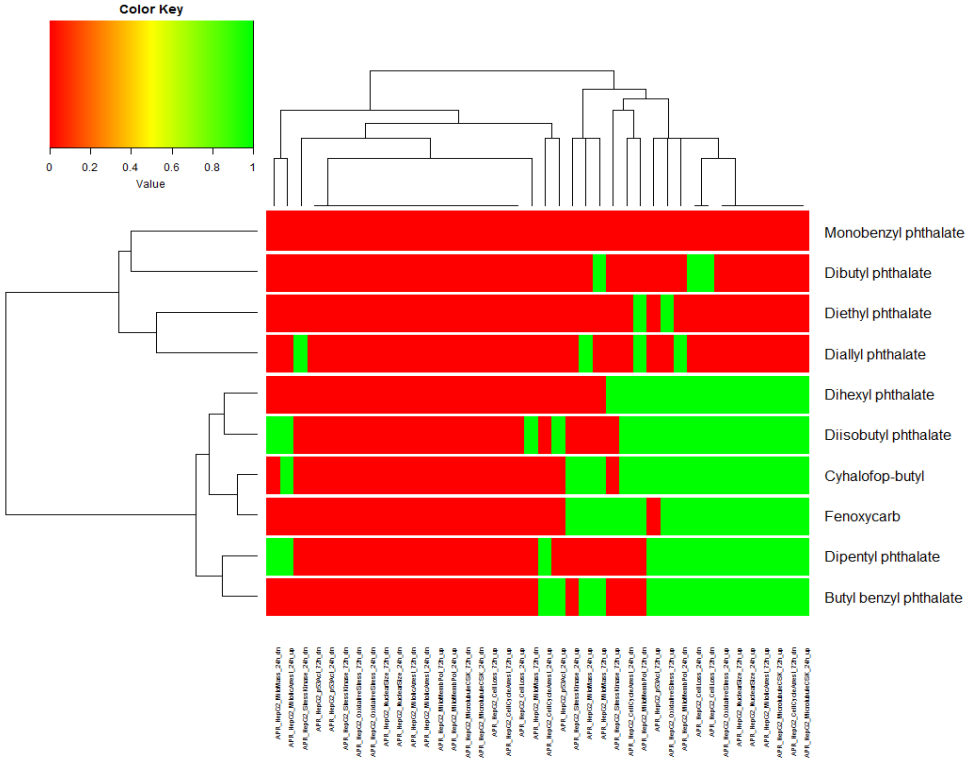


Endpoint	Baseline Prediction	Structure + Pchem Prediction
Body Weight	.78	.79
Clinical Chemistry	.27	.60
Food Consumption		
Hematology		
Kidney		
Liver		
Mortality		
Pancreas		
Prostate	0	0
Skin	.27	.21
Spleen	0	.20
Tissue NOS	0	0
Urinary Bladder	0	0

- **Adding phys-chem to similarity search overturns incorrect predictions for 2 endpoints**
- **Improves many others**

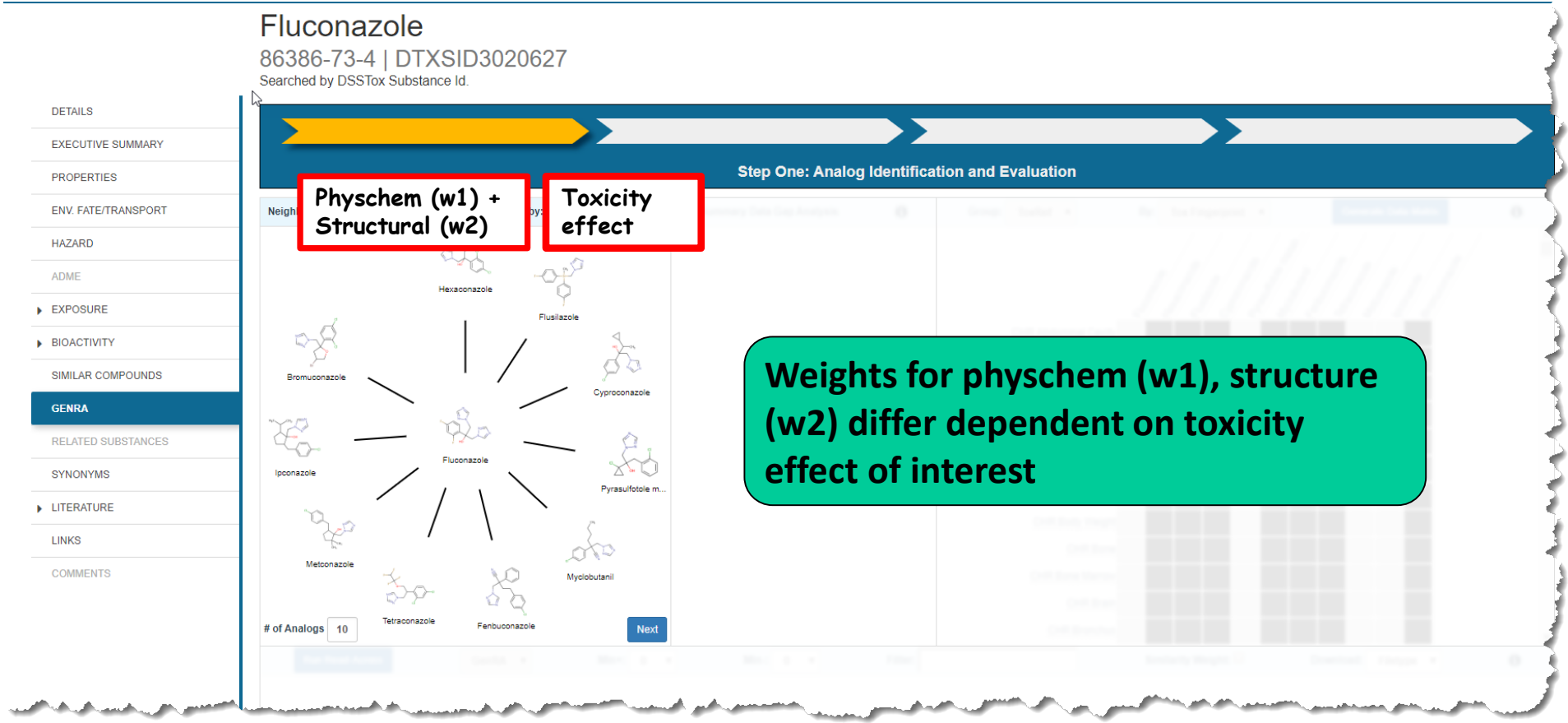
Case Study: Butyl Benzyl Phthalate

Approach 2: Search Expansion



- Are the non phthalate analogues plausible from a biological similarity context?
- Heatmap of ToxCast bioactivity profiler from one (Apredica) technology
- From a qualitative perspective – these non phthalates exhibit similarity wrt their bioactivity profile to the target and other source phthalates

“Search expansion” in practice

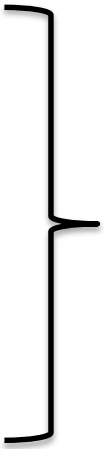


Refinements to the GenRA approach

- Transitioning GenRA from binary predictions to 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

Refinements to the GenRA approach: Acute toxicity

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

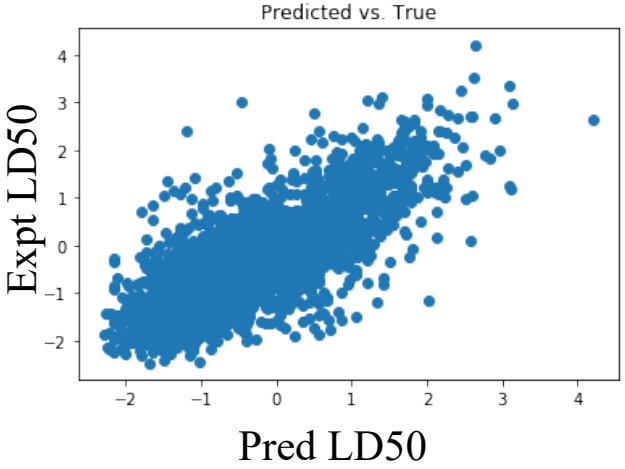
11,992 chemicals

16,209 LD50 values

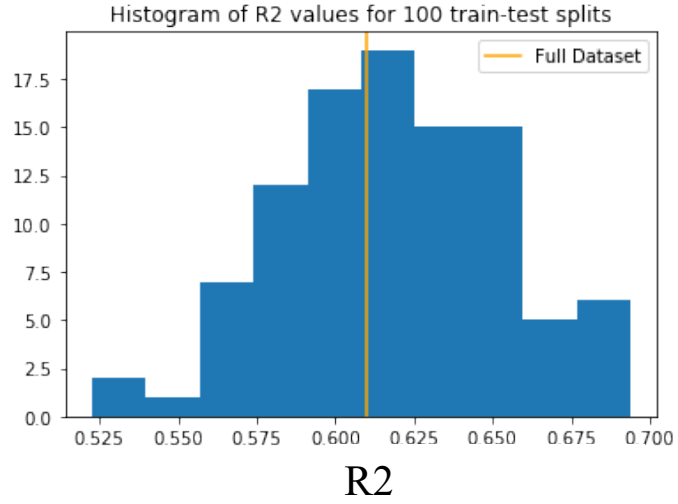
Karmaus et al, 2018; Kleinstreuer et al., 2018

Refinements to the GenRA approach: Acute toxicity

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



- R^2 = 0.61
- RMSE = 0.58
- A few outliers, but not too extreme
- Residuals clustered around zero with no obvious patterns



- 75-25 train-test splits
- R^2 values range from 0.52 to 0.69
- GenRA performs strongly and robustly on this acute tox data set.

Conclusions

- Current workflows for developing category/analogue approaches follows a series of steps
- There are many similarities between them – a harmonised version has been proposed
- There are many sources of uncertainty and proposals to address these for read-across to be more routinely accepted
- Many read-across tools exist that align to the workflow steps
- To move towards quantifying uncertainties we need to consider different approaches to structuring read-across – that will perform objective measures of performance to be determined
- GenRA has been used to illustrate some of the possibilities

Future Directions

- Include..
- Capturing other contexts of similarity such as reactivity, metabolism information and quantifying the impact on read-across performance
- Moving from quantitative predictions using classical toxicity data such as LD50 acute rodent oral toxicity to bioactivity data from HTS assays such as those generated within ToxCast/Tox21 or HTTr (Benchmark Dose/Concentration)

Acknowledgements

Imran Shah – US EPA

George Helman – US EPA

Tony Williams – US EPA

Rusty Thomas – US EPA

Jason Lambert – US EPA

Lucy Lizarraga – US EPA

Katie Paul Friedman – US EPA

References

Data Quality

- Conrad JW, Jr, Becker RA. 2011. Environ Health Perspect. 119: a508–a509.
- https://arasp.americanchemistry.com/Data-Quality-Evaluation.pdf
- https://eurl-ecvam.jrc.ec.europa.eu/about-ecvam/archive-publications/toxrtool
- Samuel GO, et al 2016 Environ Int. 92-93:630-46.

References

Guidance and examples

- **OECD, 2014:** http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=ENV/JM/MONO(2014)4&doclanguage=en

- **ECETOC TR 116:** http://www.ecetoc.org/publication/tr-116-category-approaches-read-across-qsar/

References

Frameworks for identifying analogues:

- Wu S et al 2010. Regul Toxicol Pharmacol. 56(1):67-81.
- Patlewicz G et al 2013 Regul Toxicol Pharmacol. 67(1):1-12.

References

Frameworks for assessing read-across:

- Blackburn K, Stuard SB. 2014 Regul Toxicol Pharmacol. 68(3):353-62.
- Patlewicz G et al 2014 ALTEX. 2014;31(4):387-96. Patlewicz G et al 2015 Regul Toxicol Pharmacol. 72(1):117-33.
- Schultz TW et al 2015 Regul Toxicol Pharmacol. 72(3):586-601.
- ECHA RAAF https://echa.europa.eu/documents/10162/13628/raaf_en.pdf
- Ball N et al 2016 ALTEX. 33(2):149-66.

References

New approaches in read-across

- Low Y et al 2013 Chem Res Toxicol. 26(8):1199-208.
- Shah I et al 2016 Regul Toxicol Pharmacol. 2016 79:12-24.
- Zhu H et al 2016 ALTEX. 33(2):167-82.
- Patlewicz et al 2017 Comp Toxicol
- Patlewicz et al 2018 Comp Toxicol
- Helman et al 2018 Comp Toxicol