A framework to build scientific confidence in read-across results

Grace Patlewicz, PhD

National Center for Computational Toxicology (NCCT)

US Environmental Protection Agency (US EPA)

RTP, NC

patlewicz.grace@epa.gov

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)
- Reaction domain – organic chemistry reaction mechanisms that characterise electrophilic chemicals
- GenRA – Generalised Read-across

Talk Objectives

Understanding:

- Workflow for category/analogue approaches
- Importance of the decision context
- Current read-across software tools – where within the category workflow they add most value
- Uncertainty assessment
- Future directions towards quantifying read-across performance and its associated uncertainties

Workflow for category/analogue approach

1. Decision context
2. Data gap analysis
3. Overarching hypothesis
4. Analogue identification
5. Analogue evaluation
   - Data gap filling
6. Uncertainty assessment

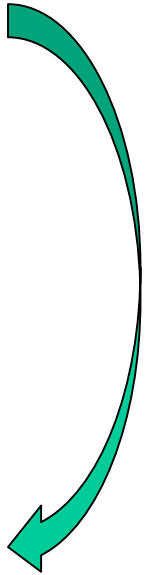
Workflow for category/analogue approach

1. Decision context
2. Data gap analysis
3. Overarching hypothesis
4. Analogue identification
5. Analogue evaluation
   - Data gap filling
6. Uncertainty assessment

Read-Across Tools

1. Decision context

- Prioritisation e.g. PMN
- Screening level hazard assessment
- Risk Assessment e.g. PPRTV



- Different decision contexts will dictate the level of uncertainty that can be tolerated

Workflow for category/analogue approach

1. Decision context
2. Data gap analysis
3. Overarching hypothesis
4. Analogue identification
5. Analogue evaluation
   - Data gap filling
6. Uncertainty assessment

Read-across Tools – An Illustrative List

Tool	OECD Toolbox	ToxMatch	AMBIT	ToxRead
Analogue identification	X	X	X	X
Analogue Evaluation	X	X	X To an extent by other predictive tools available	X For Ames & BCF
Data gap analysis	X Data matrix viewable		X Data matrix can be exported	
Availability	Free	Free	Free	Free

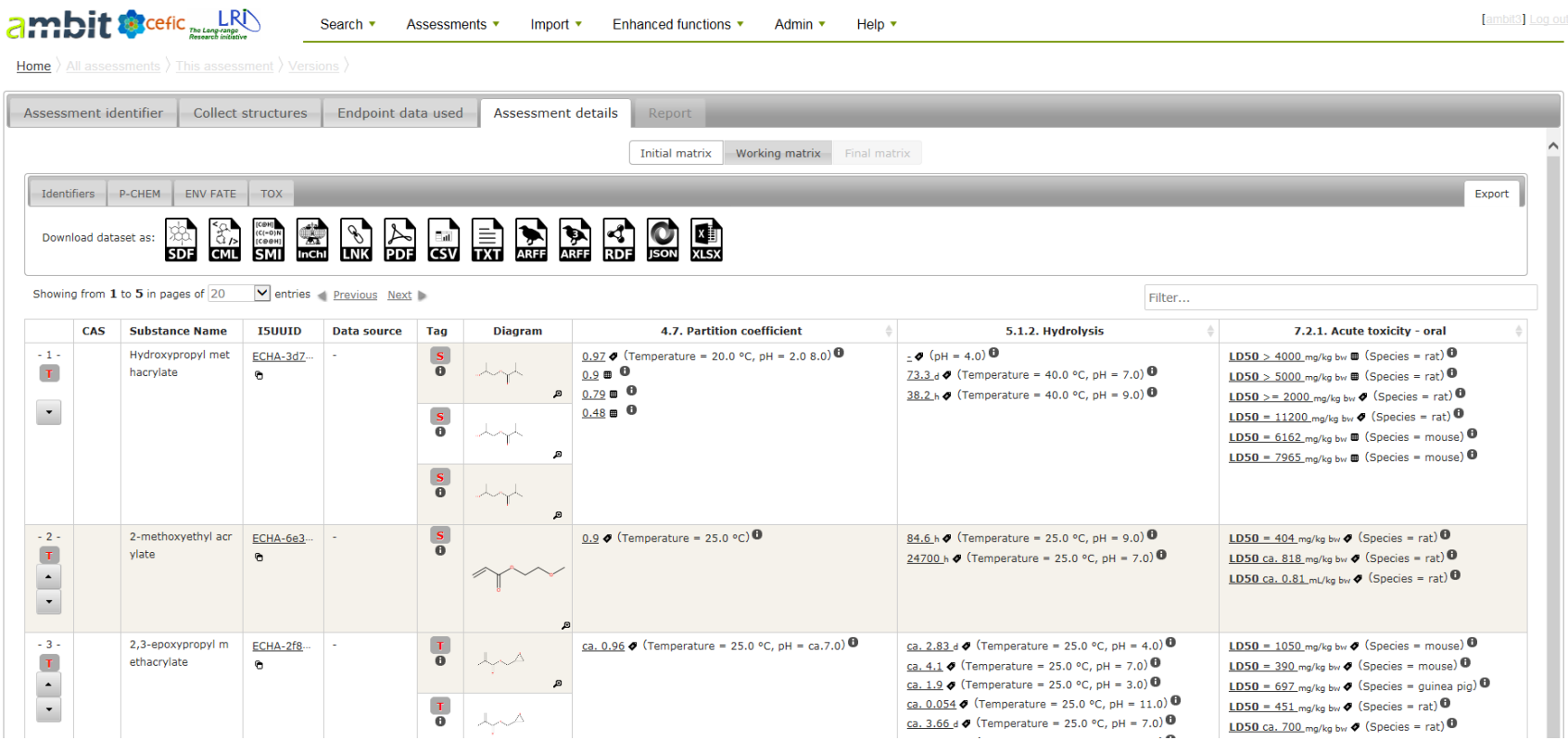
2. Data gap analysis

- Evaluating the completeness of the data matrix to identify specific data gaps for a target substance
- Depends on access to high quality study data
  - Study quality can be assessed using frameworks such as that proposed in Klimisch et al 1997
  - ToxRTool is a software tool that can facilitate such an assessment

2. Data gap analysis

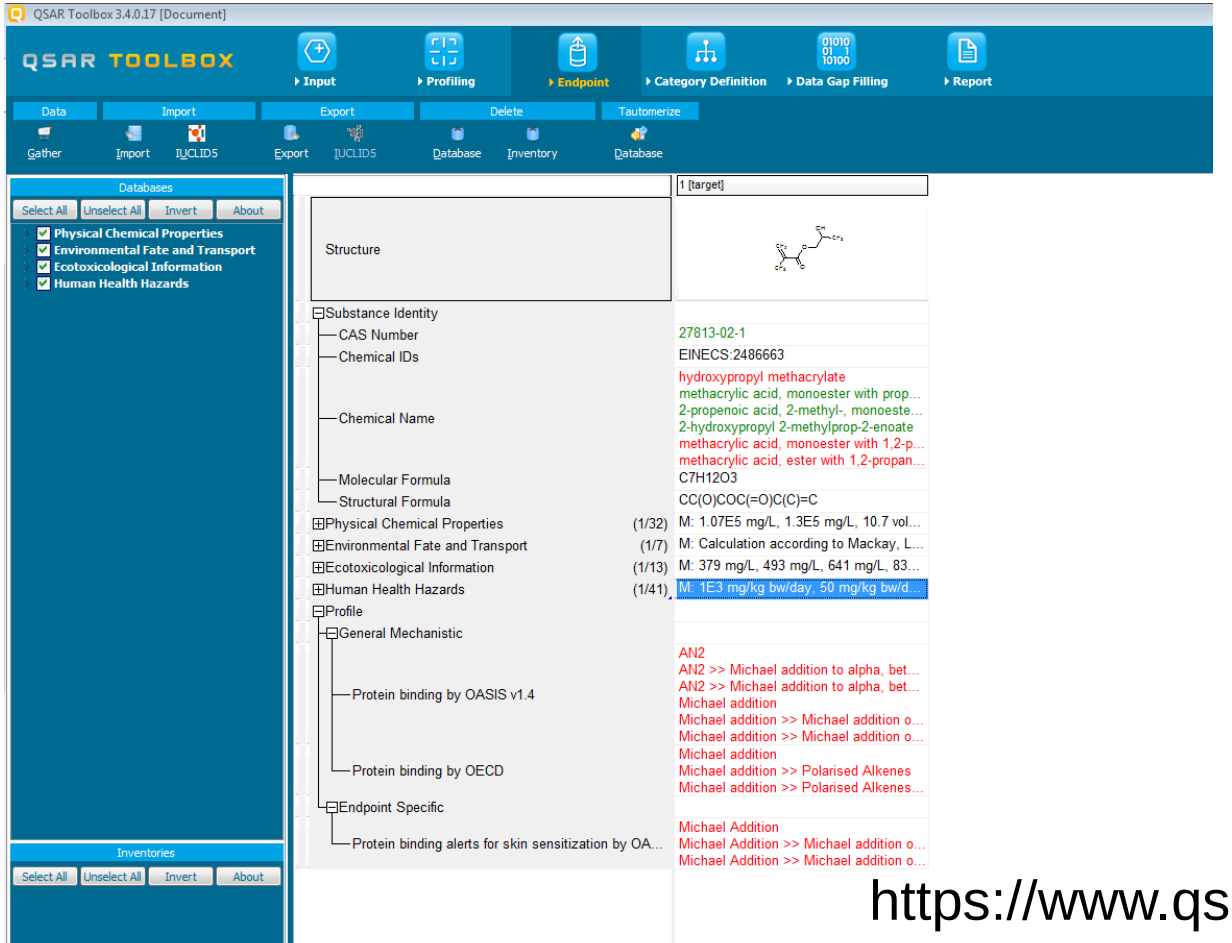
- Read-across tools that allow data gaps to be quickly identified for the target chemical include:
  - AMBIT
  - OECD Toolbox

Data matrix: AMBIT



http://cefic-lri.org/lri_toolbox/ambit/

Data matrix: OECD Toolbox



https://www.qsartoolbox.org/

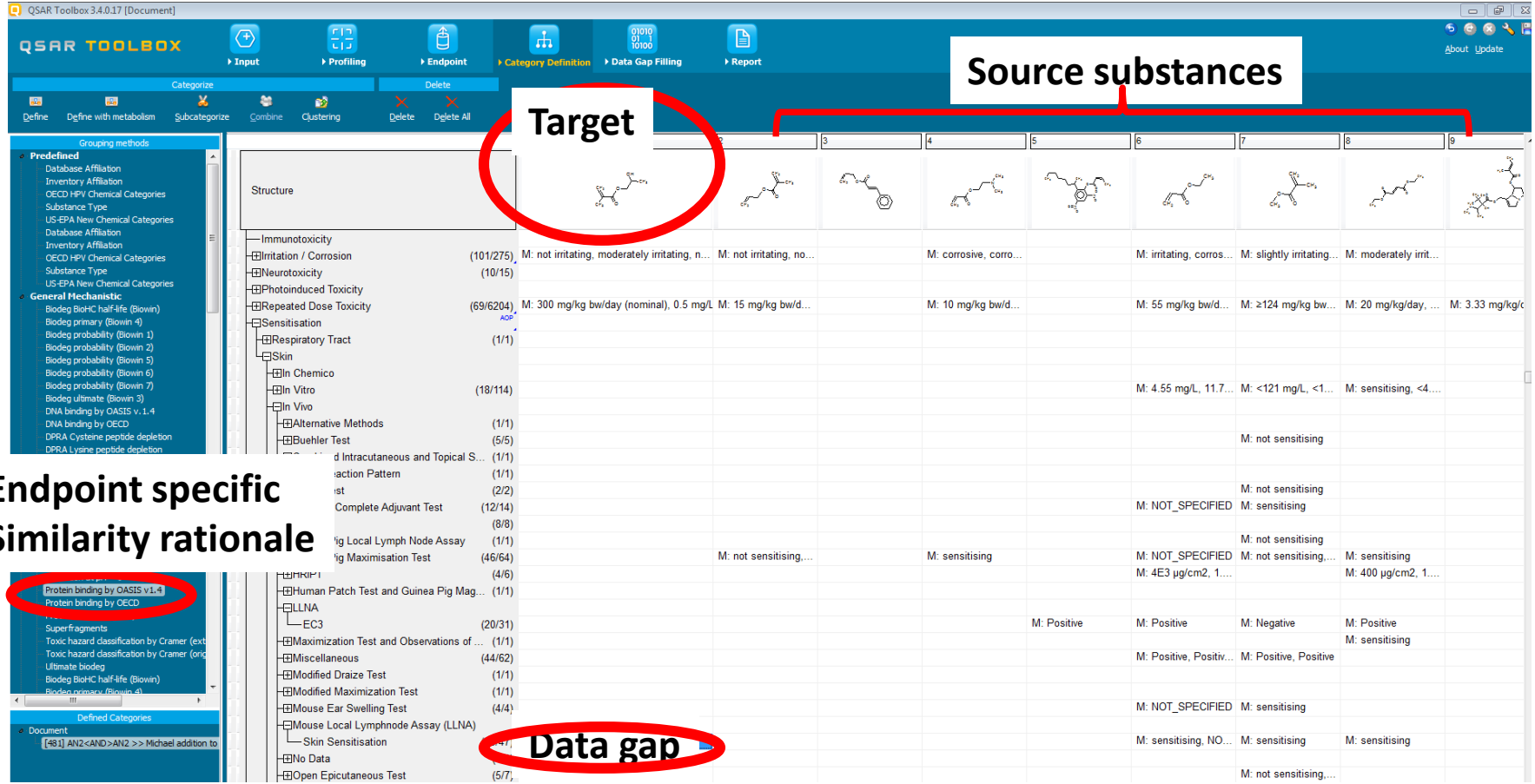
Steps 3 to 5 of the workflow

- Read-across tools that assist in identifying similar analogues and justifying their similarity for the endpoint of interest include:
  - OECD Toolbox
  - ToxMatch
  - ToxRead

Analogue identification and evaluation: OECD Toolbox

- Define an endpoint specific category to predict e.g. skin sensitisation potential for a target chemical
- Overarching similarity rationale = same protein binding alerts
- Data matrix is updated to reflect target and potential source analogues

Analogue identification and evaluation



Endpoint specific
Similarity rationale

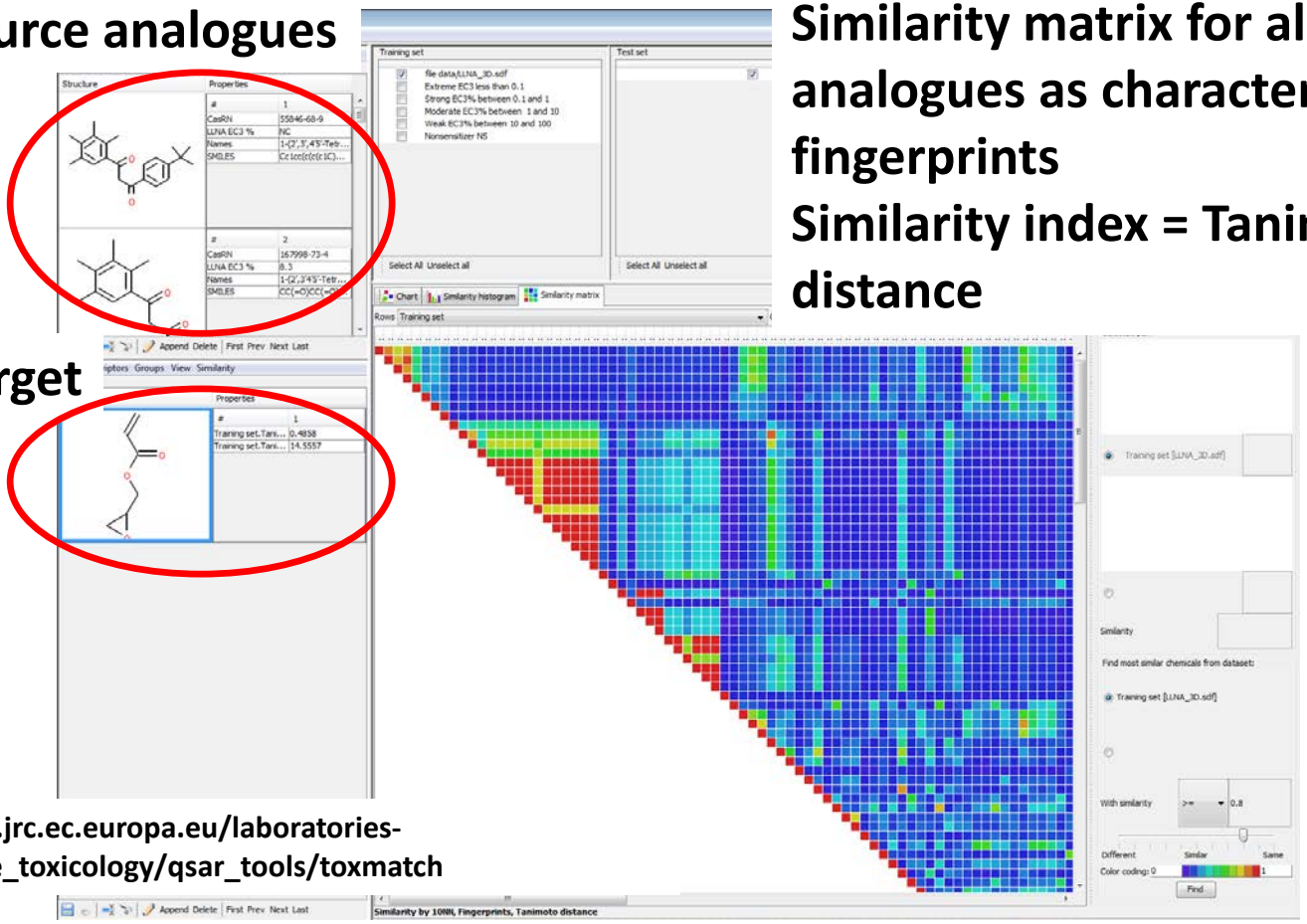
Analogue identification and evaluation: Toxmatch

- Identify similar analogues on the basis of fingerprints from a predefined dataset e.g. skin sensitisation
- Filter analogues on the basis of a similarity index threshold

Toxmatch

Source analogues

Target

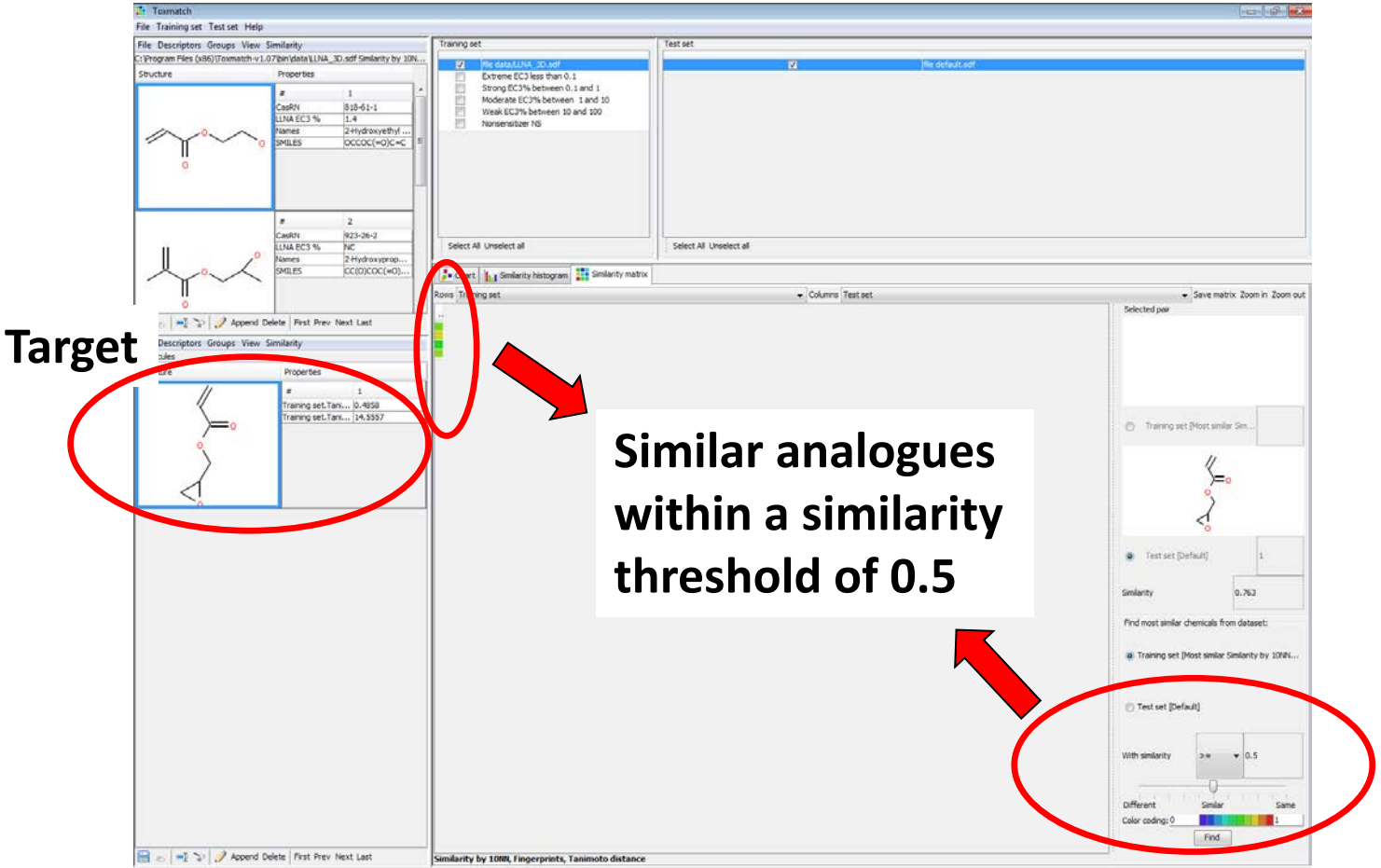


Similarity matrix for all source analogues as characterised by fingerprints

Similarity index = Tanimoto distance

https://eurl-ecvam.jrc.ec.europa.eu/laboratories-research/predictive_toxicology/qsar_tools/toxmatch

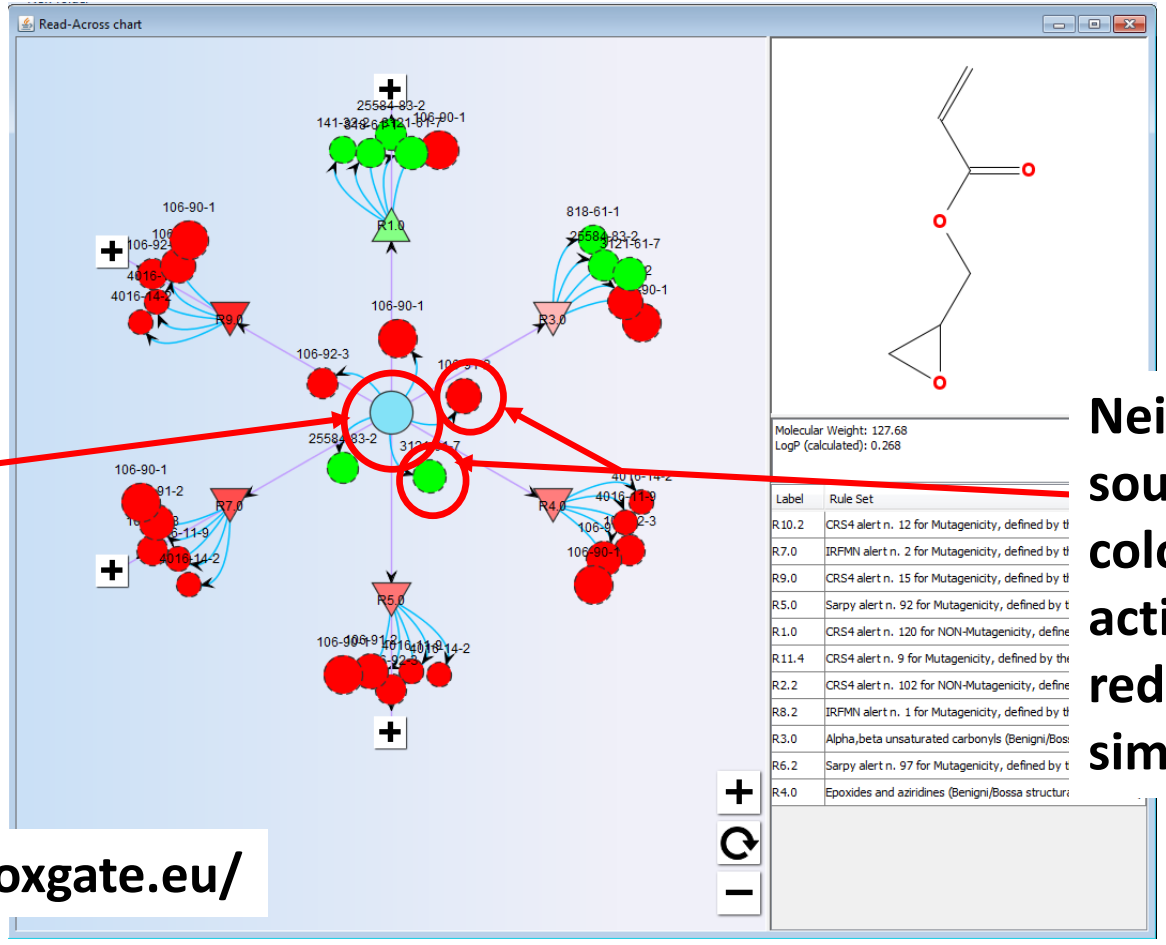
Toxmatch


Target
Similar analogues within a similarity threshold of 0.5

Analogue identification and evaluation: ToxRead

- Identify similar analogues on the basis of structural similarity and structural alerts
- Endpoints covered are mutagenicity and bioconcentration potential
- User defines number of source analogues

ToxRead



Target

Neighbouring source analogues, colour coded by activity (positive = red) and by similarity index

http://www.toxgate.eu/

6. Uncertainty assessment

- A number of publications exist that can guide the construction and assessment of categories and use of read-across
  - Guidance and examples (OECD, 2014; ECHA, 2008; ECETOC TR 116, 2012;)
  - Frameworks for identifying analogues e.g. Wu et al, 2010, Patlewicz et al, 2013
  - Frameworks for assessing read-across (Blackburn and Stuart, 2014, Patlewicz et al, 2015; Patlewicz et al, 2015; ECHA – RAAF, 2015; Schultz et al, 2015; Ball et al, 2016)

See references list for full citations

6. Sources of uncertainty

- Analogue or category approach? (# 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

6. Sources of uncertainty (cont'd)

- 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
- Presence vs absence of toxicity
- Toxicokinetics

Strategies to evaluate and address uncertainties - addressing dissimilarities

- Evaluating whether structural differences of the source analogue may impact the toxicity relative to the target substance
- Are there specific structural alerts identified for the structural features that are not common between the target and source analogues?
  - e.g. Use of systems such as the OECD Toolbox, Derek Nexus can be helpful in identifying specific structural alerts

Strategies to evaluate and address uncertainties - addressing dissimilarities (cont'd)

- Do the structural differences translate to significant differences to the metabolic pathway between source and target analogue that could result in differences in toxicity? e.g. Use of the OECD Toolbox's metabolic simulators or METEOR may prove helpful in exploring the metabolic pathways and their differences
- Do the structural differences result in significant differences to the physicochemical properties that could impart differences in bioavailability? e.g. Estimation of LogKow and MW can provide useful insights into potential differences in bioavailability

Strategies to evaluate and address uncertainties – toxicokinetics and metabolism

- Toxicokinetics – including Metabolism
  - Underlying rationale presumes a metabolic transformation e.g. Source analogue => Target
  - Assumption is that this transformation is rapid and complete
  - What sort of practical approaches can be applied to demonstrate that such transformation occurs?

Strategies to evaluate and address uncertainties – toxicokinetics and metabolism (cont'd)

- Predict likely metabolite(s) using in silico tools
  - e.g. OECD Toolbox, Meteor Nexus, MetaPrint 2D, TIMES, Catalogic
- Assessing metabolism through one or another experimental systems.
  - E.g. precision-cut tissue slices, subcellular fractions such as the microsomal fraction, primary cells (immortalized, in suspension, monolayers in culture), cell lines (continuous, liver-derived etc.)

Read-across performance

- Uncertainty that can be tolerated depends on the decision context
- However read-across acceptance relies on a subjective expert assessment
- Uncertainty assessment is qualitative in nature
- There is no objective measure of read-across performance
- But there are efforts in progress

(NB: previous presentation)

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 assays (bio)
ToxCast
574 Apical outcomes (tox)
ToxRefDB

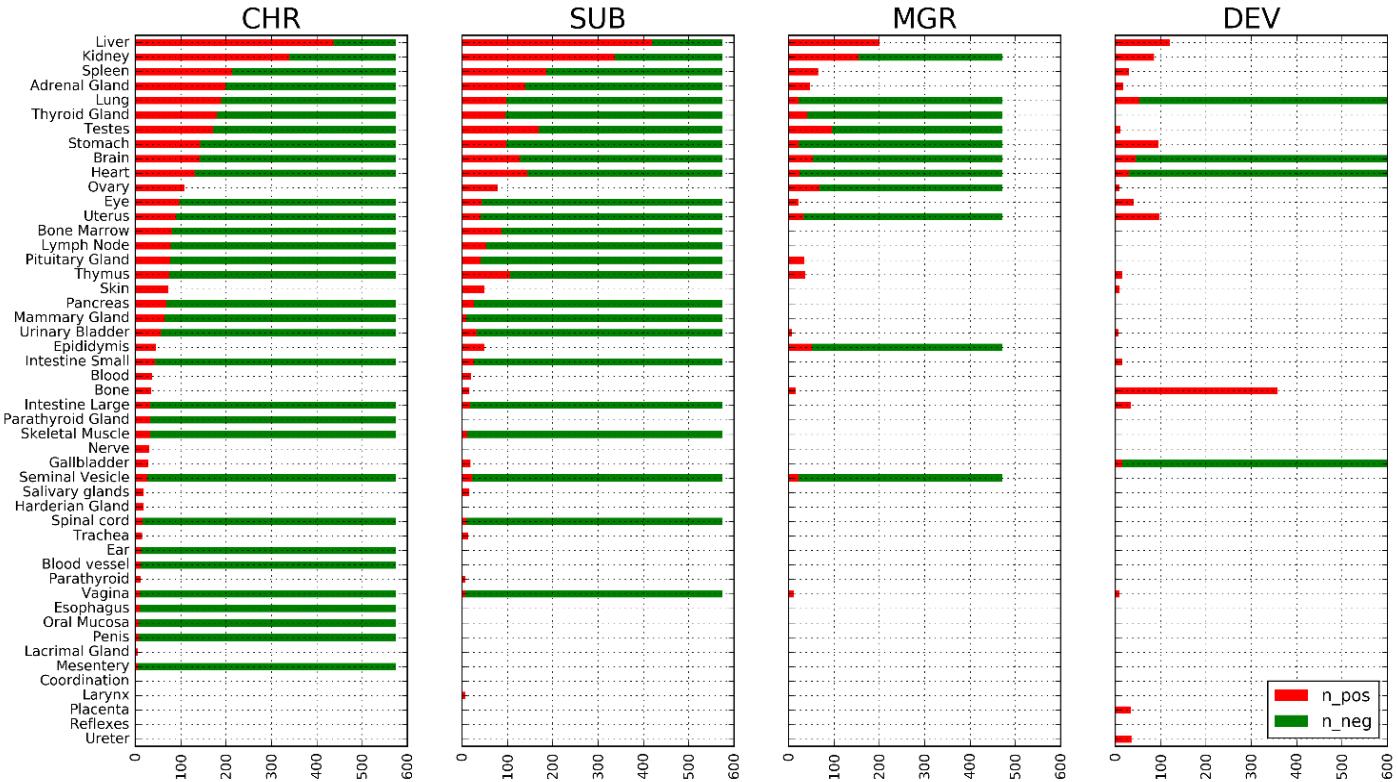
II. Define Local neighborhoods

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

III. GenRA

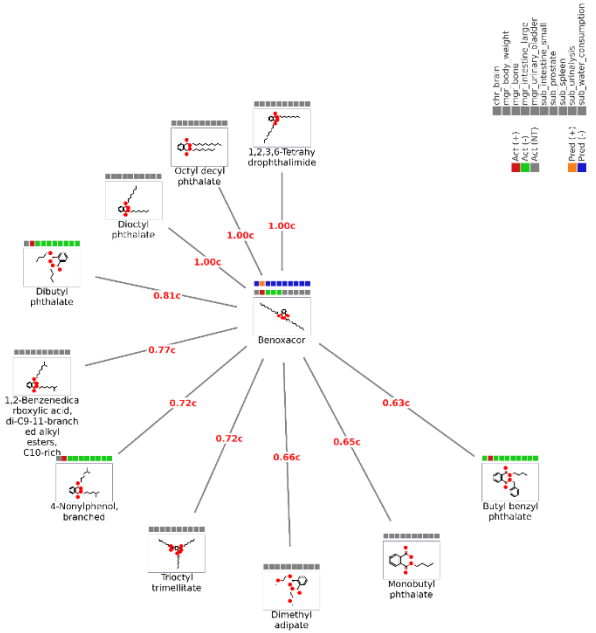
Use GenRA to predict apical outcomes in local neighbor hoods
Evaluate impact descriptors (chm, bio, bc) on prediction
Quantify uncertainty

GenRA - Toxicity Data from ToxRefDB

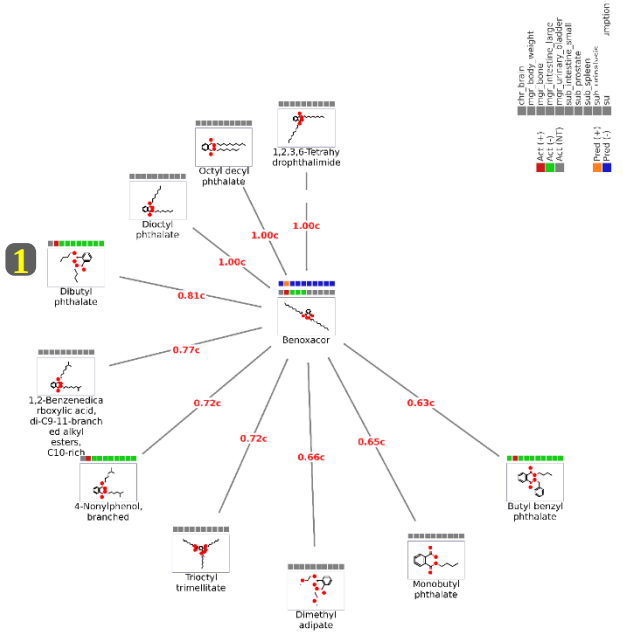


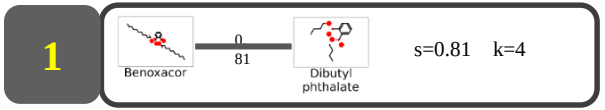
GenRA – performance in each cluster

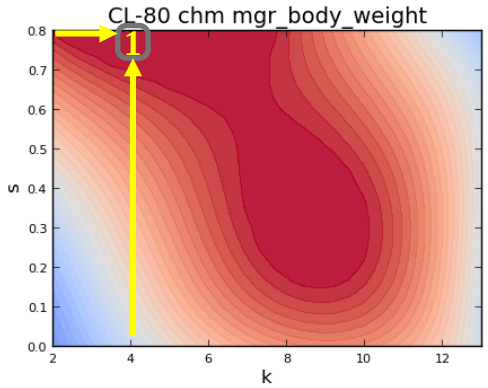
- Use GenRA to predict the similarity weighted toxicity scores for each
  - Toxicity type (β)
  - Descriptor ={chm,bio,bc} (α)
  - No. of nearest neighbours (k)
  - Similarity score threshold (s_{ij}^{α})
- Calculate performance by comparing predicted y^{tox} and true x^{tox} for all chemicals using area under ROC curve (AUC)
- Results: {cluster, α, β, k, s, AUC}



GenRA - Analyzing local neighborhood of a chemical







GenRA – Insights and Next steps

- Bioactivity descriptors were often found to be more predictive of *in vivo* toxicity outcomes
- The approach enabled a performance baseline for read-across predictions of specific study outcomes to be established but was still context dependent on the endpoint and the chemical
- Next steps:
- Use of other chemical descriptor sets that encode more expert knowledge of SARs
- Incorporating TK information

Conclusions

- Current workflow for developing category/analogue approaches follows a series of steps
- Decision context is a key consideration as this will drive the level of uncertainty that can be tolerated
- There are many sources of uncertainty and proposals to address these
- To move towards quantifying uncertainties we need to consider different approaches to structuring read-across
- An example is provided to illustrate some of the possibilities

Talk Objectives

Understanding:

- Workflow for category/analogue approaches ☑
- Importance of the decision context ☑
- Current read-across software tools – where within the category workflow they add most value ☑
- Uncertainty assessment ☑
- Future directions towards quantifying read-across performance and its associated uncertainties ☑

Acknowledgements

Imran Shah – US EPA

Rusty Thomas – US EPA

Nicholas Ball – The Dow Chemical Company

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.