



Applying *In Silico* Approaches to Chemical Safety in Regulatory Programs: Genotoxic Impurities of EFSA Pesticides

Bologna SITOX
Chihae Yang,
James Rathman, Aleksandra Mostrag, Oliver Sacher,
Mark Cronin, Elena Fioravanzo

Feb 11, 2020



Motivation: Guidance Document

GUIDANCE



Adopted July 22, 2016

doi: 10.2903/j.efsa.2016.4549

Guidance on the establishment of the residue definition for dietary risk assessment

EFSA Panel on Plant Protection Products and their Residues (PPR)

Abstract

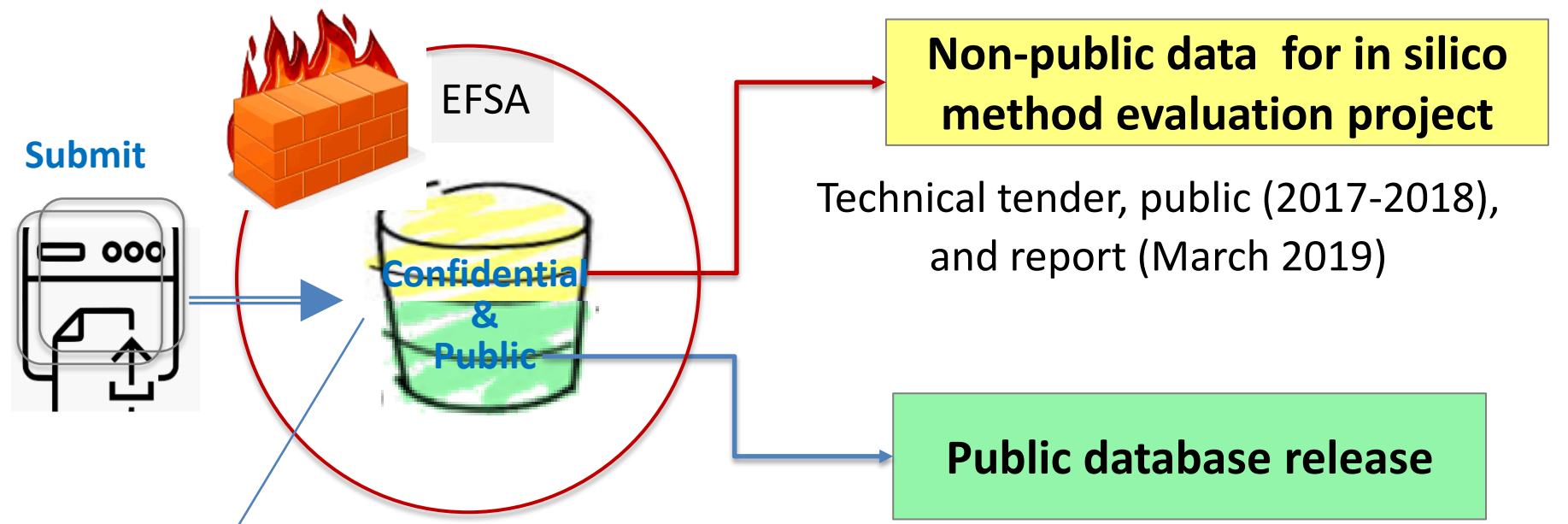
EFSA has asked the Panel on Plant Protection Products and their Residues to prepare guidance on the establishment of the residue definition for dietary risk assessment. The residue definition for risk assessment is used by risk assessors to evaluate the potential risk of dietary intake of residues resulting from the application of a pesticide. This document guides the complex process of identifying the pertinent residue components that should be considered for dietary risk assessments of chemical active substances. Specifically, the document provides directions for determining the metabolites that require hazard identification and characterisation using scientific tools and methods ((quantitative) structure-activity relationship ((Q)SAR), read-across, threshold of toxicological concern (TTC)) and available data in combination, and for developing an appropriate testing strategy for these

The Residue Definition for Dietary Risk Assessment

Need: evaluation of applicability of existing (Q)SAR models and Read Across approaches for prediction of genotoxicity of pesticides and their metabolites



Data Flow Across European Food Safety Authority Example: Genetic Toxicity Database



<https://doi.org/10.2903/sp.efsa.2017.EN-1229>

Compilation of a database, specific for the pesticide active substance and their metabolites, comprising the main genotoxicity endpoints 17 May 2017

Evaluation of *In Silico* Methods to Address Genotoxic Impurities of EFSA Pesticides

EXTERNAL SCIENTIFIC REPORT



APPROVED: 12 March 2019

doi:10.2903/sp.efsa.2019.EN-1598

Evaluation of the applicability of existing (Q)SAR models for predicting the genotoxicity of pesticides and similarity analysis related with genotoxicity of pesticides for facilitating of grouping and read across

Romualdo Benigni*, Chiara Laura Battistelli**, Cecilia Bossa**, Alessandro Giuliani**, Elena Fioravanzo**, Arianna Bassan**, Mojca Fuart Gatnik**, James Rathman****, Chihae Yang****and Olga Tcheremenskaia**

*Alpha-Pretox (www.alpha-pretox.com), Rome, Italy

**Istituto Superiore di Sanità (ISS), Environmental and Health Department, Rome, Italy

***S-IN Soluzioni Informatiche Srl, Vicenza, Italy

****Altamira, LLC, USA

Abstract

Project Objectives Overview

I

Critical review of existing QSAR models

Partners: Istituto Superiore di Sanità, Alpha Pre-Tox

II

Thorough description of prediction results by QSAR

Modelers: 5 Commercial and 3 public providers
Organized & analyzed by S-IN, ISS, Alpha Pre-Tox, MN-AM (Altamira)

III

Critical review of existing Read-Across & tools

Partners: Istituto Superiore di Sanità, Alpha Pre-Tox

IV

Thorough description of Read-Across results and evaluation by case studies

Partners: Alpha Pre-Tox, MN-AM (Altamira), Istituto Superiore di Sanità



Outline

Data Overview

QSAR

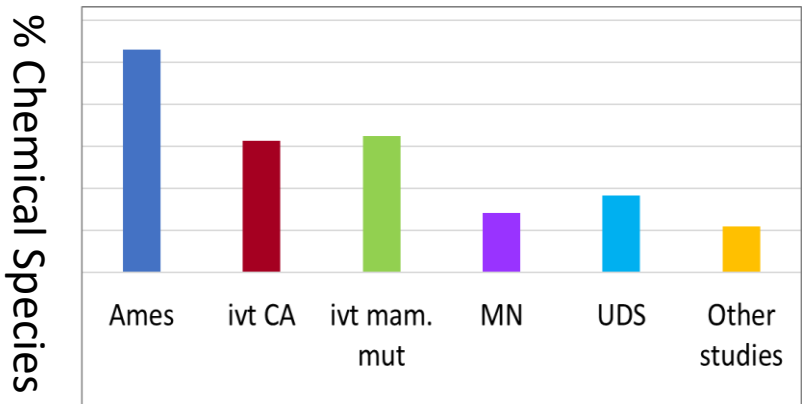
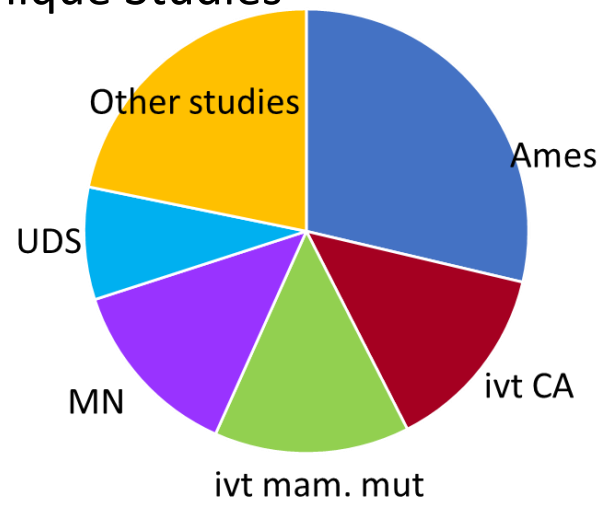
Read-Across

Conclusion



Genetic toxicity Database of Pesticides & Metabolites

By Unique Studies



- A total of 1109 chemical species
 - 380 parents and 1035 metabolites and impurities
- Data for 23 genetic toxicity test types in 5,561 unique studies (24,721 tests)
- Data sources in EFSA DAR
 - EC opinions (4,177 tests for 241 chemicals)
 - EFSA opinions (12,083 tests for 723 chemicals)

Other studies include CA, DNA damage & repair, dominant lethal, cell transformations assays, etc.



Genotoxicity Profiles

% Positives in Genotoxicity Data

Assay	%POS (Compounds)	%POS (Studies)
Reverse bacterial mutagenesis (Ames)	7.6	8.9
In vitro chromosome aberration	26.7	26.7
Micronucleus	18.8	10.5
In vitro mammalian mutagenesis	13.8	13.0



Outline

Data Overview

QSAR

Read-Across

Lessons Learned



Blindfolded model provider

QSAR Experimental Design

- **5 Commercial and 3 Public Providers**
 - **Statistical:** ACD/Labs, Lazar, Leadscope, Lhasa (Sarah), MultiCASE, Vega
 - **Rule-based:** Lhasa (DEREK), Toxtree, Vega
 - **Combined (or hybrid) methods:** ChemTunes·ToxGPS® (MN-AM)
- **Endpoints**
 - Bacterial reverse mutagenesis (Ames) – 19 models (11 QSAR, 8 Expert Rules)
 - In vitro chromosome aberration – 7 models
 - In vivo micronucleus – 6 models (restrictions due to biology)



Profile Data – Bacterial Reverse Mutagenesis

- **Experimental data**

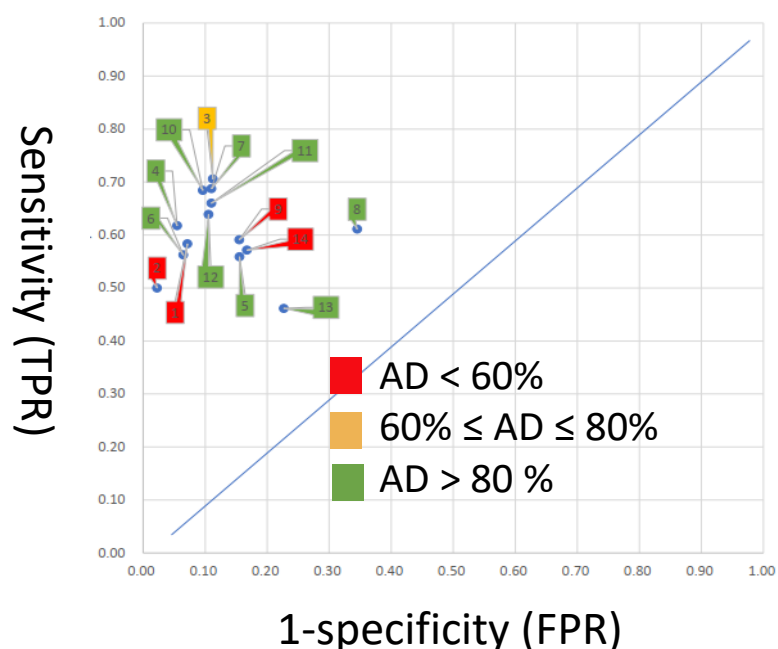
- Negatives: 879 (92%)
- Positives: 39 (4%)
- No Calls: 38 (4%)
- All data points: 956
- Total [POS+NEG]: 918

- **Prediction profile**

- % Sensitivity range:
 - min: 28.6%
 - max: 68.7%
- % Specificity range:
 - min: 65.5%
 - max: 99.2%
- Best performer
 - 68.8% sensitivity / 88.8 % specificity / N=839
- Extreme case
 - 30.8 % sensitivity / 99.2 % specificity / N=393



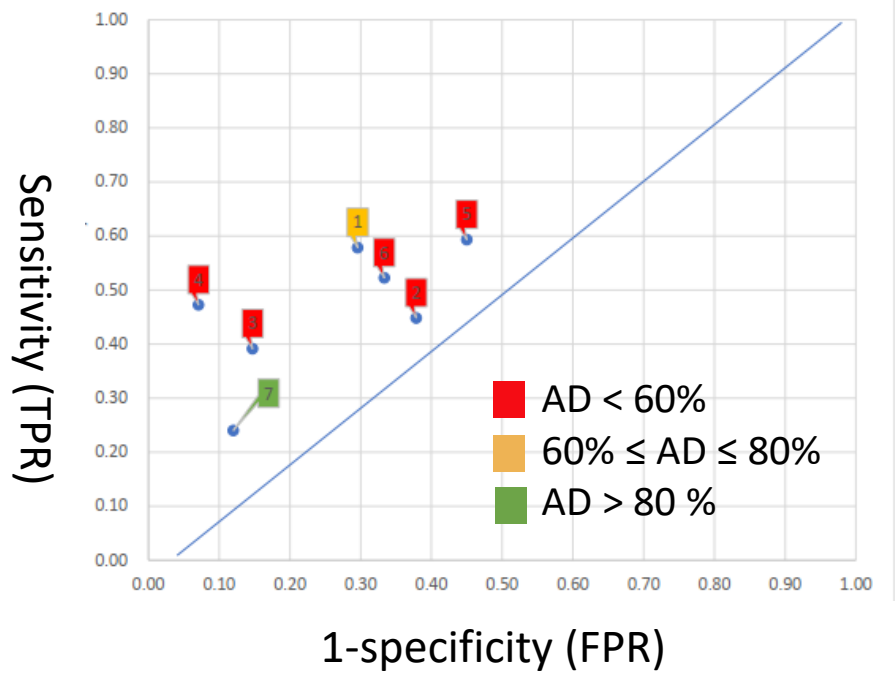
Bacterial Reverse Mutagenesis



- OECD Guideline 471
 - TA 100, 98, 1535, 1537, WP2 and WP2urvA (unless TA 102)
- Potential issues in the EFSA test dataset
 - Different biology interpretations
 - Many conflicting studies
 - Older studies
 - Only 4.7 % positive



In Vitro Chromosome Aberration



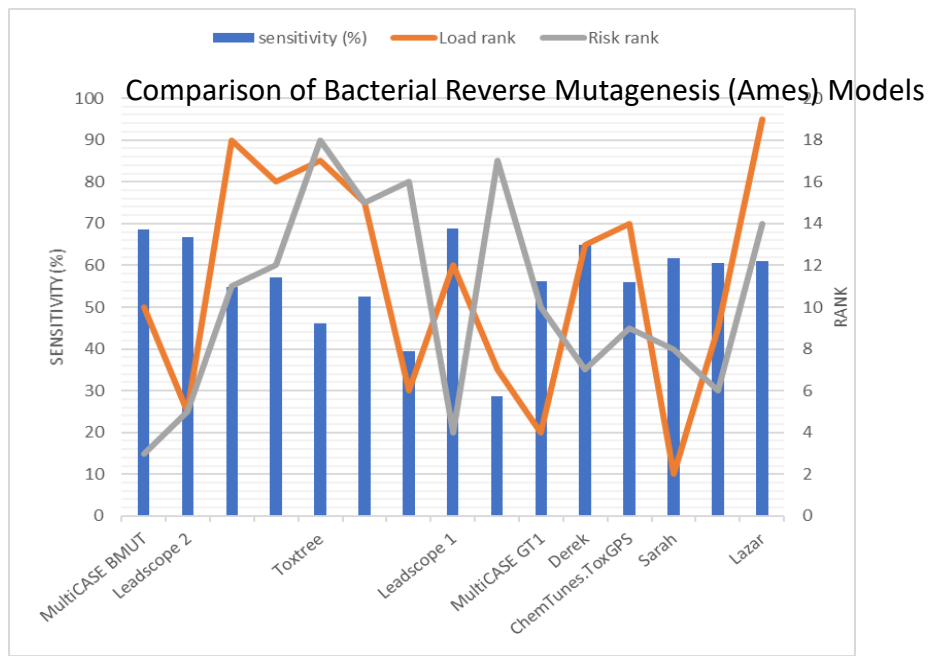
- Older experimental data are less reliable.
- Complex biology, e.g.,
 - Pulsed and continuous experiments or cell line types.
 - Project partners did not consider biology

- NTP study protocol does not conform to OECD473 guideline, hence difficult to be used as negative compounds.



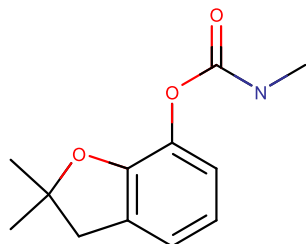
Selection of the Reliable Models:

- “2-best” models defined by highest sensitivity in this study.
- “Selected 7-models” for further work
 - sensitivity $\geq 55\%$
 - specificity $\geq 85\%$
 - % compounds in domain of applicability $\geq 80\%$
- Ranking
 - Assay load (FP)
 - Risk (FN)



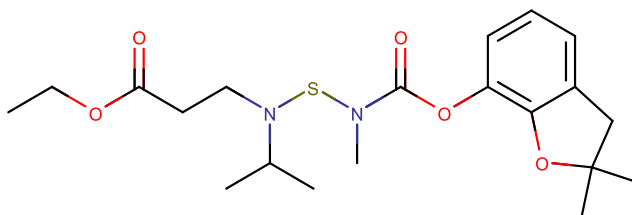
Common False Negative (Benigni's call based on data)

Carbofuran (CMS-257)



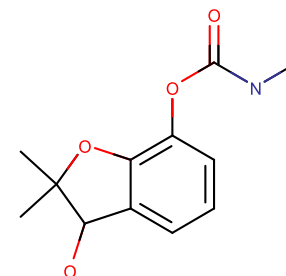
Negative (1985), Positive (1983)

Benfuracarb (CMS-8934)



Negative (1982)

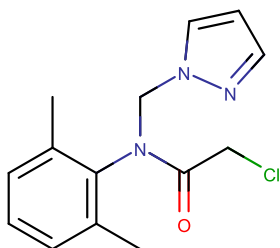
3-OHCarbofuran (CMS-7634)



Negative (2003), Positive (2002)

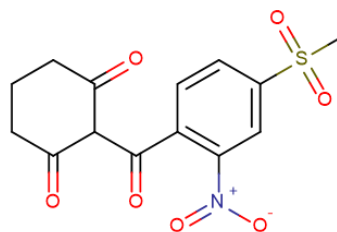
Common False Positive (Benigni's call based on data)

Metazachlor (CMS-58046)



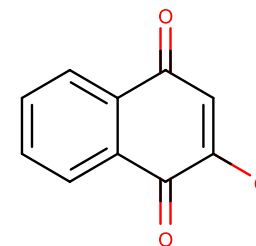
4 Negative (1994, 2001-2006)

Mesotrione (CMS-6762)



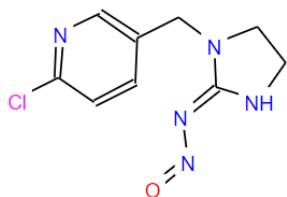
Positive (1998), 2 Negative (1993, 2013)

Henna Dye (CMS-4356)



1 Negative (2003)

In Silico Combinations for Outcome



In database: Yes
 CHEMTUNES ID: CMS-203819
 Name: 1-[(6-CHLORO-3-PYRIDINYL)METHYL]-,OXOHYDRAZONE
 Registry number(s):
 # studies in CHEMTUNES: 0

In this example, the Global model alone gives a false positive prediction, but combination of the Global and Organohalide models gives the correct prediction (negative).

Predictions for endpoint: Bacterial Reverse Mutagenicity

Show Details: [MoA models](#) [Chemotype alerts](#) [Nearest neighbors](#)

NEGATIVE

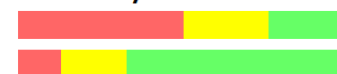
Probability (POSITIVE) = [0.331, 0.414]



MoA Models

Model name	Probability (POSITIVE)
Global	[0.519, 0.779]
Ames Organohalide 2015/1	[0.132, 0.332]

Probability Bar

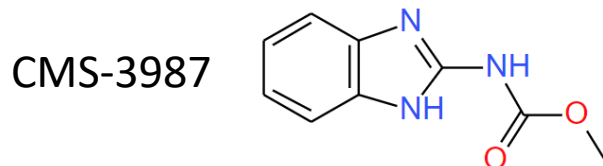




Weight of Evidence Approach

Combination of results from multiple models

- outperforms any single model
- broadens the knowledgebase

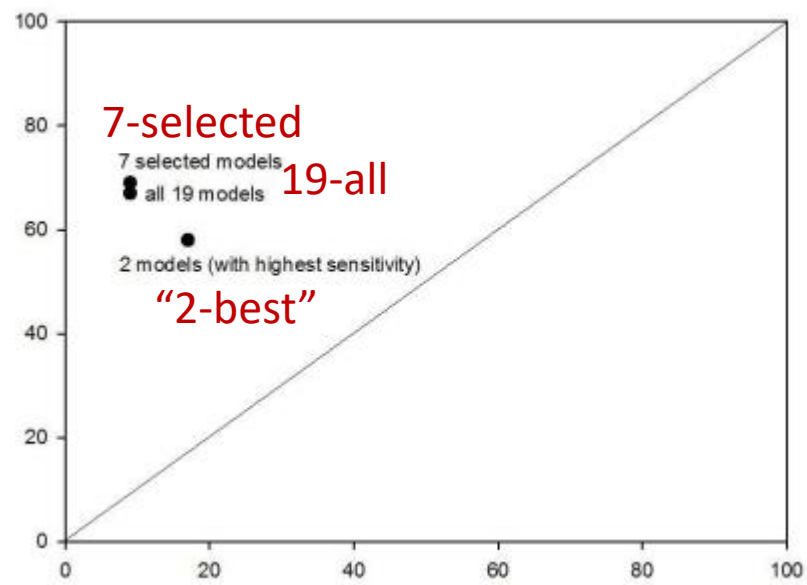


Rathman, J.F., Yang, C., Zhou, H. "Dempster-Shafer theory for combining in silico evidence and estimating uncertainty in chemical risk assessment", *Computational Toxicology* 6, 16-31 (2018)



Combination of Models by Weight of Evidence Approach

- Both sensitivity and specificity of Ames models increased when models were combined by decision theory approach.
 - “Selected 7-models” outperforms all models combined or individually.



Models Combined	% Sensitivity	% Specificity
“2 Best”*	58	83
7 selected	69	91
All 19	67	91



Observations

- Results support the hypothesis that WoE predictions obtained from multiple models may be better than for any individual model alone.
- The extremely low number of positives (4%, 39 out of 918) in this test set makes it difficult to make any strong conclusions when comparing different models.



Outline

Data Overview

QSAR

Read-Across

Lessons Learned

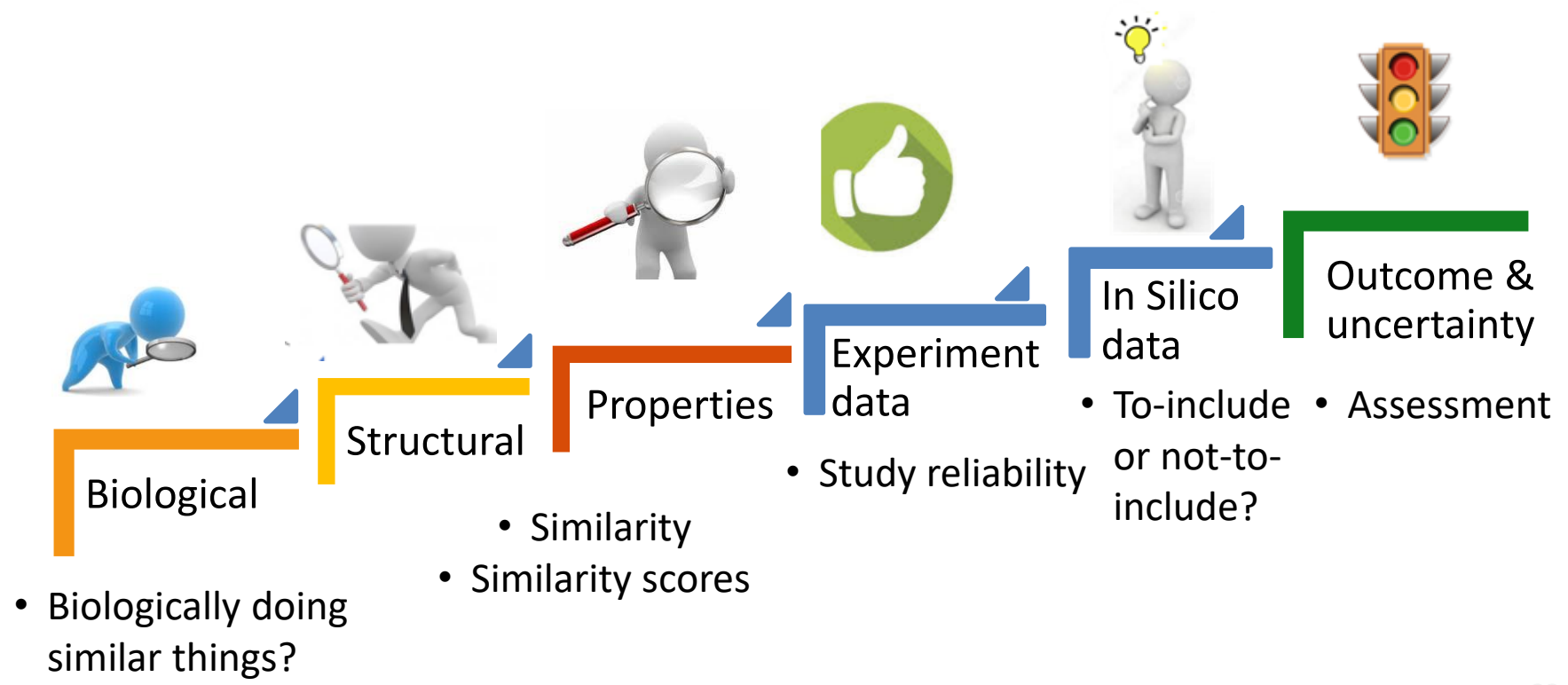


Objective 4: Use Case Studies of Read-Across

- **General strategic assumption**
 - Submissions provide experimental data for active parents;
 - Data for metabolites are often not available.
- **Objective 4 assumption**
 - For simplicity, prediction of a metabolite is based on an active (parent) (1:1 read-across)



Workflow of the Read-Across Process





General Read-Across Strategies: 3 Common Scenarios

- **Scenario 1: Mechanistic knowledge is available**
 - Hypothesis is known (e.g., AOP pathway known)
- **Scenario 2: Mechanistic knowledge lacking**
 - Generation of hypothesis
 - Analogue similarity assessment is critical
 - Biological & Chemical (Structure and Properties)
- **Scenario 3: Chemical reactivity or biotransformation can influence the RA process**
 - Metabolic reactions may produce compounds very different from parent



Biological Similarity – Pesticidal MoA

- Pesticidal MOA was applied to group chemicals
- **Pesticidal MOA chemical groups are NOT related to genetic toxicity mechanisms**
 - Only used to cluster chemical space

CMS-202453 1	544	CMS-203150 1	585	HPPD inhibitors 1 HPPD inhibitors 2 Picolinic acid herb. 3 7
CMS-203200 1	623	CMS-203221 1	637	
CMS-203792 1	744	CMS-203794 1	746	
CMS-203879 1	830	CMS-204042 1	960	

Navigation: <<< 11 / 19 >>> <<< 3 / 3 >>>

ChemoTyper (<https://chemotyper.org/>) and ToxPrint chemotypes (<https://toxprint.org/>) are public tools.



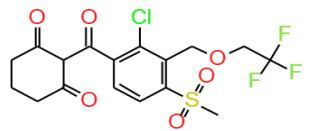
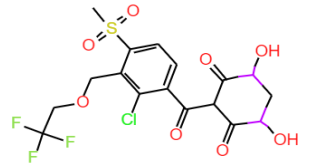
Chemical Measures of Similarity

- Chemical structures
- Chemical reactivity (biotransformation reactivity)
- Physicochemical properties



Similarities by Structure Fingerprints - Parent & Metabolite Pair

RDKit MolFingerprint similarity (Tanimoto coefficient)

<p>CMS-11340</p> 	<p>0.92</p>
<p>0.78</p>	<p>CMS-203792</p> 

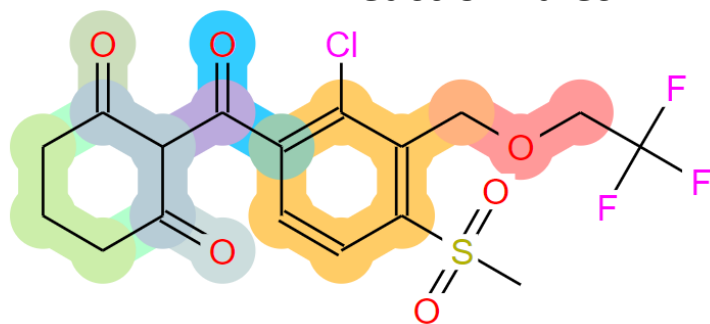
- Commonly used fingerprints: RDKit, PubChem, etc.
- **Mechanistic fingerprints**
 - ToxPrint chemotypes used to fingerprint molecules
 - Tanimoto coefficient provides pairwise similarities

ToxPrint Fingerprint similarity (Tanimoto coefficient)

Metabolic Reactivity Similarity between Parent and Metabolite

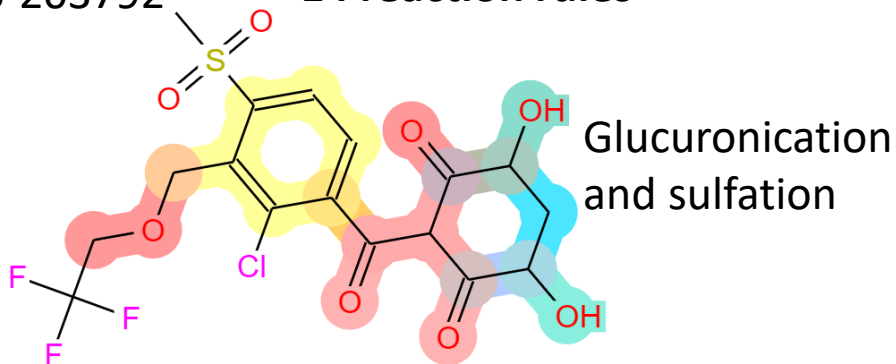
CMS-11340

11 reaction rules



CMS-203792

14 reaction rules


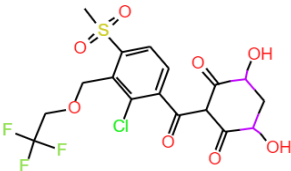


- **Metabolic reactivity chemotypes**

- The presence of a particular rule indicates the presence of a metabolic reaction site.
- Publicly available set of metabolic rules (SyGMA rules)

Similarities by Structure & Metabolic Reactivity

Liver BioPath Fingerprint similarity
(Tanimoto coefficient)

<p>CMS-11340</p> 	<p>0.5</p>
<p>0.79</p>	<p>CMS-203792</p> 

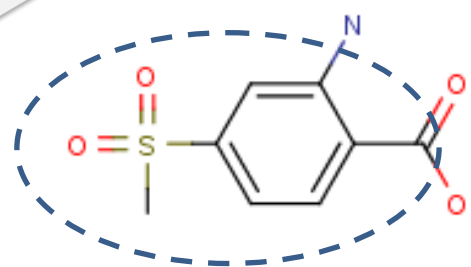
Liver BioPath Fingerprint similarity
(Metabolic reactivity similarity = M/P)

- Metabolic Reactivity Similarity = $\frac{M}{P}$
 - P is the total number of metabolic reaction sites in the parent
 - M is the number of sites common to both parent and metabolite
- Tanimoto Coeff

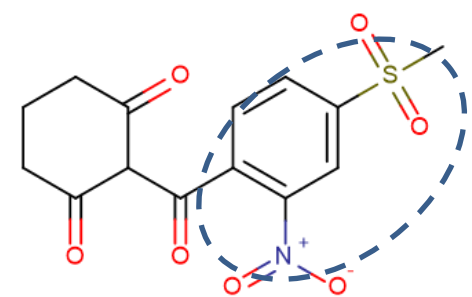
$$\text{Tanimoto Coeff} = \frac{\text{Common in P \& M}}{P + M - \text{Common}}$$



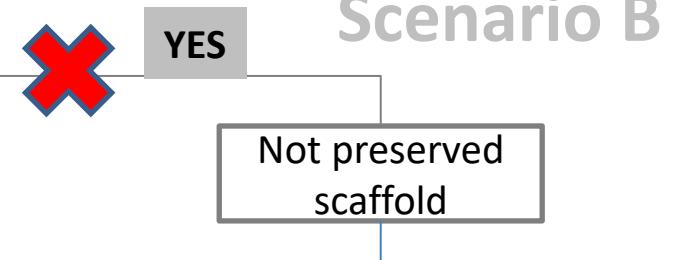
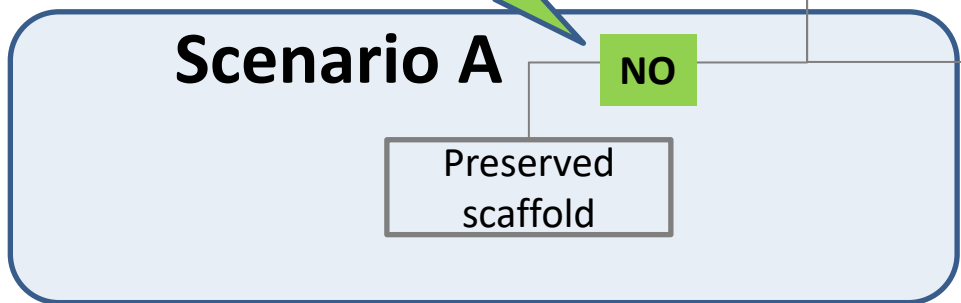
Reading Metabolite Data from Parent is Feasible when the MOA Scaffold Is Preserved.



Select an MOA group



Was the MOA scaffold changed during metabolic reactions?





Chemical Measures of Similarity

- **Molecular and physicochemical properties**
 - Properties calculated using CORINA Symphony Community Edition (public)
 - Available through MN-AM web service, UE EPA Dashboard, COSMOS NG

number of H-bond acceptors

complexity

Number of H-bond donors

topological polar surface area

molecular weight

polarizability

molecular complexity

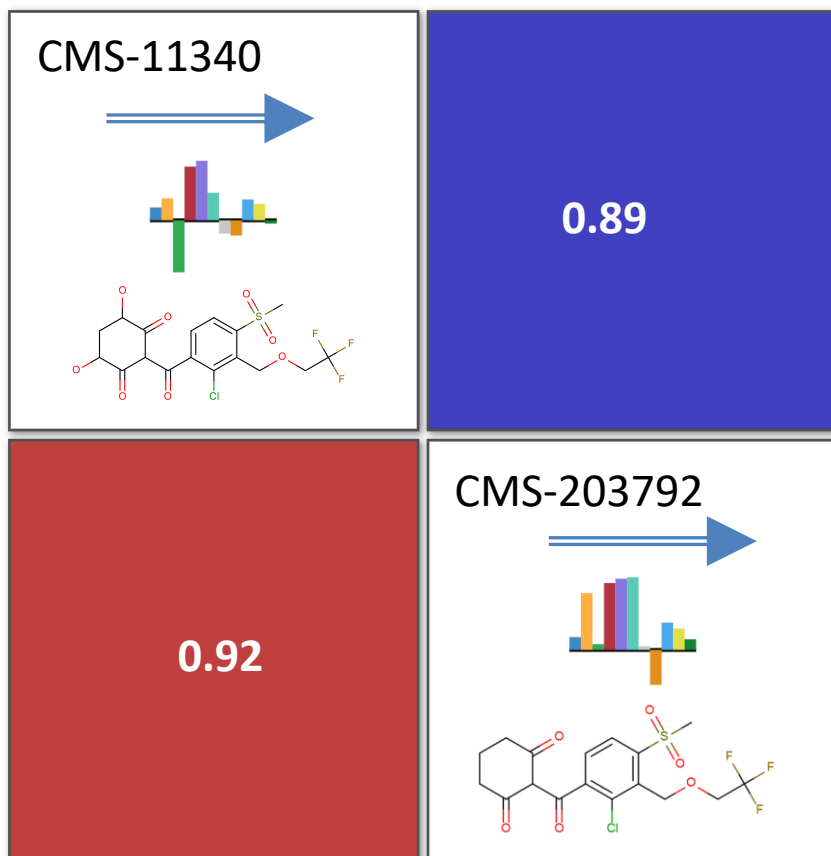
water solubility

McGowan volume

octanol/water partition coefficient

Similarities Between Profiles of Properties

Pearson similarity



Euclidean distance

- Standardized property values for a given compound
- Pearson correlation-based similarity measure

of Rotational bonds

of Hydrogen bond acceptors

of Hydrogen bond donors

of Hydrogen bond donors

Complexity

Topological Polar Surface Area

LogS

LogP

McGowan

Polarizability

Diameter



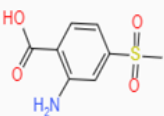
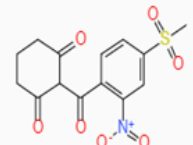


Experimental Study Reliability

Five factors considered when rating a study:

- 1) **OECD** or equivalent guideline and deviation
- 2) **GLP** compliance
- 3) **Study design - test system** (species, strains, cell lines, metabolic activation)
- 4) **Study design – test conditions** (concentration, dose levels and ranges, number of duplicates, repeated experiments)
- 5) **Study design - control** information

Reliability Score	Description	Example
1.0	Meets all five requirements as well as the number of revertant counts at a given conc. level are available along with the precipitation and cytotoxicity.	If we reviewed the conc. / dose level data from study records that satisfy all , then the study reliability would be 1.0.
0.95	Meets all five requirements, but no detailed conc. level reading.	Although followed OECD guideline, only calls were available.
0.85	Studies either missing records or not conducted and at least one deficiency in the five aspects.	The deviation included the highest concentration did not cover the full range recommended.
0.70	Studies either missing records or not conducted and at least two deficiencies in the five aspects.	If the test strains lacked WP2 or TA102, but the outcome was negative.
0.55	Studies either missing records or not conducted and more than two deficiencies in the five aspects.	If the OECD guidelines had deviation of the test system, and only one test was done with control data not providing details.



Compound Summary		Target (Parent)	Analogue (Metabolite)
	CMS ID	CMS-202453	CMS-6762
			
Fingerprints Similarity			
	RDKit MolFingerprint		0.47
	ToxPrint Fingerprint		
	Tanimoto		0.24
	Liver BioPath Fingerprint		
	Tanimoto		0.19
Properties Similarity (Skyline Profile)			
	Skyline		
	Pearson similarity		0.73
Biological Similarity			
	Pesticidal MOA	HPPD Inhibitors	HPPD Inhibitors
Analogue Quality			0.4

Experimental Data Evaluation



		Target	Analogue
Experimental Ames Data-1	Study Design		OECD 471 Eq; Deviation (no WP2 strains, no repeats), non-GLP, Control-no data, other acceptable
	Study Outcome		POSITIVE
	Study Reliability		0.55
	reliability score: Study Reliability		
Experimental Ames Data-2	Study Design		OECD 471, No deviation (strain, dose OK), GLP, Control OK, other acceptable
	Study Outcome		NEGATIVE
	Study Reliability		0.95
	reliability score: Study Reliability		
Experimental Ames Data-3	Study Design		OECD 471, No deviation (strains, dose OK), GLP, Control-no data, other acceptable
	Study Outcome		NEGATIVE
	Study Reliability		0.80
	reliability score: Study Reliability		



Using Analogue Evidence Only

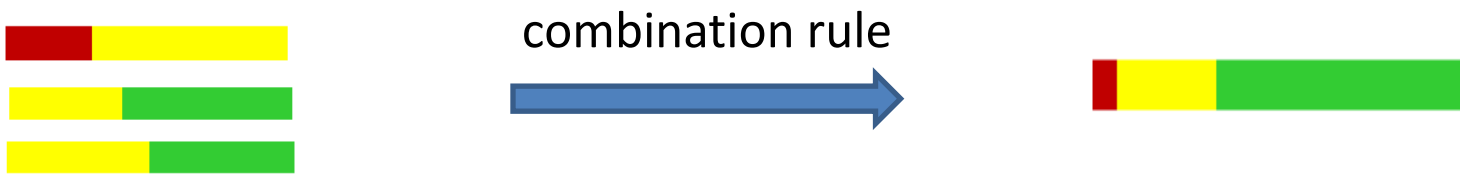
- **STEP1: Analogue Evidence**
 - Experimental outcome weighted by analogue quality and study reliability
- **STEP2: A rigorous weight of evidence combination based on decision theory**
 - Outcome and uncertainty calculated from **analogue evidence**

Experimental Study Result	Study Reliability	Analog Quality	pPOS	pNEG	Uncertainty	Probability bar
POSITIVE	0.50	0.62	0.31	0	0.69	
NEGATIVE	0.95		0	0.59	0.41	
NEGATIVE	0.80		0	0.50	0.50	



WoE Combination of Experimental Data

- **STEP3: Combine** the three experimental study results
- **STEP 4:** To obtain a **weight-of-evidence outcome** based solely on the available analogue evidence (experimental study data and analogue quality).



Analysis Outcome (Exp. only): NEGATIVE

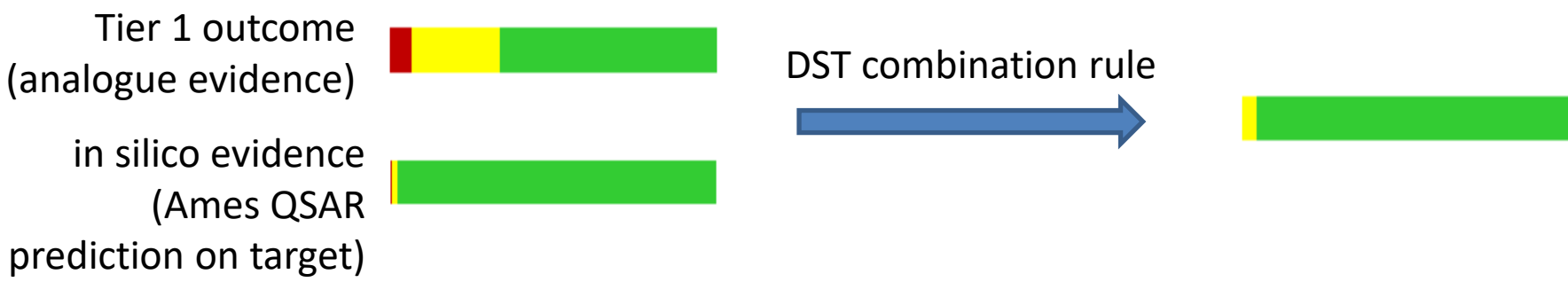
prob(POS) = 0.07 | uncertainty = 0.26

The RA prediction was correct: True Negative.



WoE Combination of Experiment + QSAR Data

- This analysis considers in silico (QSAR) prediction for the target as an additional source of evidence.



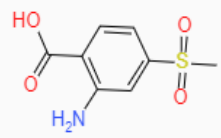
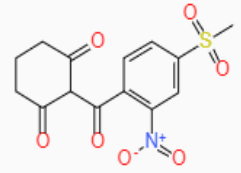


Analysis Outcome (Exp. + QSAR): NEGATIVE

prob(POS) = 0.006 | uncertainty = 0.046

Reduced the uncertainty significantly!

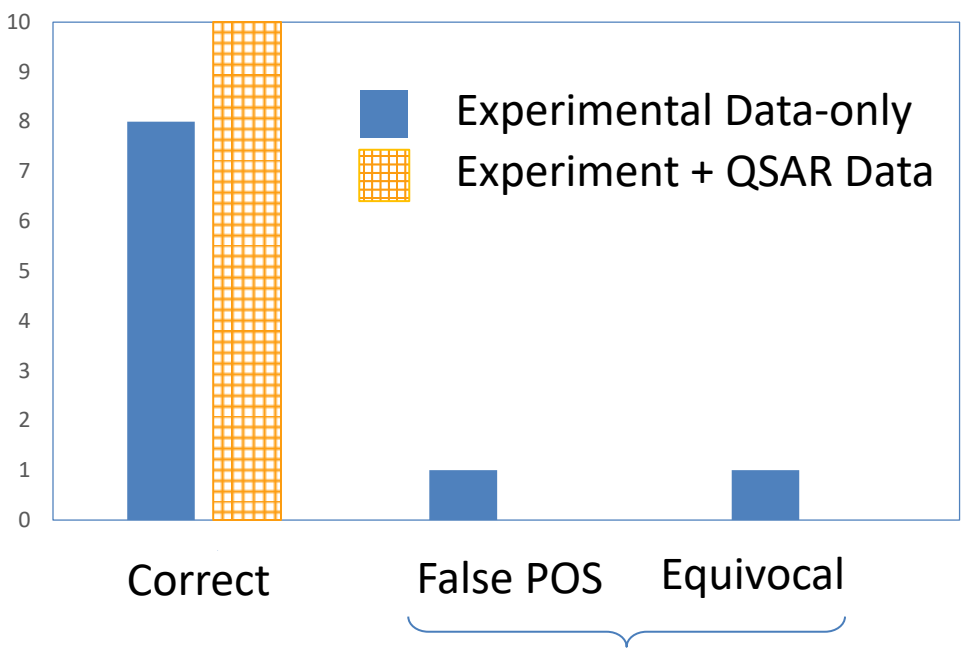


WOE Assessment Summary

Compound Summary		CMS ID	Target	Analogue
			CMS-202453 	CMS-6762 
Analogue Quality				0.4
TIER 1 (Analogue+Exp)			 0.1 - 0.5 0.39	
TIER 2 (Analogue+Exp+In silico)			 0.1 - 0.5 0.39	



RA Accuracy for Bacterial Reverse Mutagenesis



Inclusion of QSAR evidence from Ames mutagenicity models improved read-across accuracy and reduces equivocal outcomes.



Conclusion

- In silico models for Bacterial reverse mutagenesis may be a good alternative, especially when combining models from different knowledgebases.
- Read-Across of metabolite on parent requires understanding of the metabolic reactivity and retaining of the biologically meaningful scaffold.
- Use of Read-Across including a QSAR model may improve the reduction of uncertainties in the RA.



Acknowledgement

- Rositsa Serafimova, EFSA
- MN-AM (Molecular Networks and Altamira)
- S-IN
- ISS
- Aldo Benigni