

## Methods for Estimating Relative Potency Values

### Grace Patlewicz

patlewicz.grace@epa.gov Center for Computational Toxicology and Exposure (CCTE)

### March 16, 2022

Assessment Managers: Geniece M. Lehmann, PhD Krista Christensen, PhD



Office of Research and Development Center for Public Health and Environmental Assessment, Integrated Risk Information System



## **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.
- We have no conflicts of interest to disclose.



## Public Science Meeting on PCB Mixture Assessment Methods

- Introduction to EPA's human health risk assessment practices for chemical mixtures
  - Glenn Rice, U.S. EPA
- Mixtures modeling: methods considered for the assessment of PCBs
  - Laura Carlson and Jeff Gift, U.S. EPA
- Methods for estimating relative potency values
   Grace Patlewicz, U.S. EPA
- Mixture Similarity Tool (MiST)
  - Graham Glen, ICF



## **Talk Outline**

- What does Computational Toxicology encompass?
- What are approaches that can be used to fill data gaps?
- How does this help us estimate relative potency values?
- Case example using PCB Neurotoxicity Data



### **Computational (in silico) Toxicology**

- Existing information on the chemical of interest
- Predictions from (Q)SAR
- Thresholds for Toxicological Concern (TTC)
- Information from "similar" chemicals grouping/read-across
- In chemico tests
- In vitro tests
- Molecular biology, -omics
- Exposure, (bio-)kinetics



### Resources for Computational (*in silico*) Toxicology

- The <u>EPA CompTox Chemicals Dashboard</u> is just one of many existing public resources that can be used to conveniently access information from traditional and novel technologies for a large number of substances.
  - Existing information on the chemical of interest
  - Predictions from (Q)SAR
  - Thresholds for Toxicological Concern (TTC)
  - Information from "similar" chemicals grouping/read-across
  - In chemico tests
  - In vitro tests
  - Molecular biology, -omics
  - Exposure, (bio-)kinetics



### **EPA CompTox Chemicals Dashboard**

- A publicly accessible website delivering access to:
  - ~900,000 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/dashboard/



## The EPA CompTox Portal https://comptox.epa.gov/





### **CompTox Chemicals Dashboard:** Landing Page for a specific chemical

CompTox Chemicals	Dashboard Home Search - Lists - About -	Tools - Submit Comments	Se
		Welcome to the new EPA CompTox Chemicals Dashboard The new Dashboard is a complete rebuild and is replacing the CompTox Chemicals Dashboard released on July 12th 2020. This documentation can help get you started.	
	Bisphenol A 80-05-7   DTXSID702 Searched by Approved Name.	Selected a 'data-rich' substance	
Details		Wikipedia	~
Executive Summary Properties Env. Fate/Transport Hazard Safety > GHS Data ADME > IVIVE Exposure	H <sub>3</sub> C CH <sub>3</sub>	Bisphenol A (BPA) is a chemical compound and one of the simplest and best known bisphenols. It is produced by the condensation of phenol and acetone, with an estimated 4 million tonnes of produced worldwide in 2015. It is a colourless solid which is soluble in organic solvents, but poorly soluble in water (0.344 wt % at 83 °C). BPA and its derivatives have many uses, most of which are centred around plastics. Its largest single application is as a co-monomer in the Read more	d
Bioactivity Similar Compounds GenRA Related Substances Synonyms	но <b>`</b> ОН @	Intrinsic Properties  Molecular Formula: C15H1602 MOLFILE Q.FND ALL CHEMICALS  Average Mass: 228.291 g/mol Let Gotopic MASS DISTRIBUTION Monoisotopic Mass: 228.11503 g/mol	^
Literature		Structural Identifiers	~
Links		Linked Substances	~
Comments		Presence in Lists	~
		Record Information	~



Disculs and A

### **CompTox Chemicals Dashboard:** Executive Summary of 'existing' data

I ne new Dashboard is a complete rebuild and is replacing the Complex Chemicals Dashboard released on July 12th 2020. This documentation can help get you started.

	Bisphenol A 80-05-7   DTXSID7020182 Searched by Approved Name.			
Details	Executive Summary			
Executive Summary	Quantitative Risk Assessment Values	Regional Screening <sup>1</sup>		
Executive Summary	© Risvalue available ?			
Properties	No PPRTV values		Bink	
Env. Fate/Transport	S ERA RSL values available (2	Class	≡ Risk 1 ↑ Value 2 ↑ Level 1 ↑	
	🚱 Minimum RHD:0.05 mg/kg-day ( chronic, ) (2*	RFDo (mg/kg-day)	5.00e-2	
Hazard	S No RIC calculated	risk-based SSL (mg/kg soil)	THQ = 0.1 5.80	
Safety > GHS Data	VIVE POD net calculated	screening level (tap water) (ug/L)	THQ = 0.1 77.0	
	Quantitative Hazard Values	screening level (residential soil) (mg/kg soil)	THQ = 0.1 320	
ADME > IVIVE	Minimum crail PDD:0.003 mg/kg-day (Immunetexicity, crail ) (27	screening level (industrial soil) (mg/kg soil)	THQ = 0.1 4.10e+3	
Exposure	🕑 inhalasion POD values 10 mg/m3 ( subchronic, inhalasion ) (2'		THQ = 1 58.0	
	🕑 Lowest Observed Biosotivity Equivalent Level	risk-based SSL (mg/kg soil)		
Bioactivity	OVP1A1, OVP1A2, ESR1, NR118, NA, ESR1, PRARA, ESR1, ESR1	screening level (tap water) (ug/L)	THQ = 1 770	
Similar Compounds	Cancer Information	screening level (residential soil) (mg/kg soil)	THQ = 1 3.20e+3	
	🐼 No cancer alope factor	screening level (industrial soil) (mg/kg soil)	THQ = 1 4.10e+4	
GenRA	S No cancer unit risk values			
Related Substances	📀 No cancer data			
	Genotovicity Datapredicted to be clastogenic (?)			
Synonyms	Reproductive Toxicology			
Literature	Reproductive toxicity PODs available (7	PhysChem Parameters 🖲		
Links	Chronic Toxicology			
LIIKS	Chronic toxicity PODs available (?	3.32	1.64	-7.17
Comments	Subchronic Toxicology			
	Subchrolic toxicby PODs available (7	-s 0 s 10 1s logKow	-5 0 5 10 15 log(BCF)	-4 -2 0 2 log(VP)
	Calculations, basicly note an annual to	logiton	109(001)	10g(VP)
	Developmental Toxicology			
	🚱 Developmental toxicity PODs available 🖓	Point-of-Departure Plots		
	Acute Toxicology			
	Acute toxicity PODs available (?		Oral POD	
			<u> </u>	rat subchronic NOAEL rat subchronic NEL
	Subacute Toxicology		<b></b>	rat subchronic LOAEL rat subchronic LEL rat subchronic LEL
	🕑 Subsoure toxicity PODs available (?			ret short-term LOAEL
	Endocrine System		• • • •	rat reproduction NOAEL
	Endocrine Disruption Potential Significant Estrogen and Androgen Receptor activity seen. Chemical was positive in 17 ER assay(out of 21) and was positive in 9 AR assay(tested in 17) 3			rat reproduction LOAEL rat reproduction LEL rat repeat dose LOAEL
	ADME		••	ret immunotoxicity LOEL
	ADML     OHTH Cas data are available (?			rat developmental NULAEL rat developmental neurotoxicity NOAEL rat developmental neurotoxicity NEL rat developmental neurotoxicity LOAEL
				rat developmental neurotoxicity LOAEL rat developmental neurotoxicity LOAEL
	- Esta Transmart		<b>_</b>	



### **CompTox Chemicals Dashboard:** Landing Page for a specific chemical

CompTox Chemicals D	ashboard Hom	e Search <del>-</del> Lists - About -	Tools -	Submit C	Comments S	earch all data 🛛 🗸 🔍
			The ne	Welcome to the new EPA CompTox Chemicals Dashboard v Dashboard is a complete rebuild and is replacing the CompTox Chemicals Dashboard released on July 12th 2020. This documentation can help get you started.		
Details	Chemical Details	PCB 026 38444-81-4   DTXS Searched by Approved Name		In contrast, PCB 026 is 'data-po	or'	. 0
Executive Summary Properties Env. Fate/Transport	CI			Quality Control Notes SRS/ChemID matched: SRS trust index 3	^	0
Hazard Safety > GHS Data ADME > IVIVE Exposure				Intrinsic Properties Molecular Formula: C12H7Cl3 A MOL FILE Q_FNID ALL CHEMICALS Average Mass: 257.54 g/mol Mit 6010PE MASS Distribution Monoisotopic Mass: 255.961333 g/mol	^	
Bioactivity *		CI		Structural Identifiers	~	
GenRA Related Substances Synonyms	® CI			Linked Substances Presence in Lists		
Literature Links	G	-		Record Information	~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~	



### **CompTox Chemicals Dashboard:** Executive Summary of 'existing' data

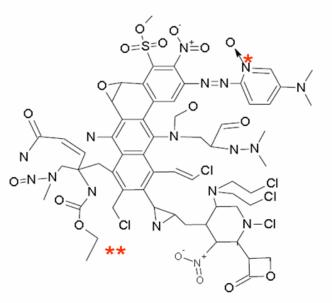
CompTox Chemica	IS Dashboard Home Search - Lists - About - Tools -	Submit Comments
	The new l	Welcome to the new EPA CompTox Chemicals Dashboard Dashboard is a complete rebuild and is replacing the CompTox Chemicals Dashboard released on July 12th 2020. This documentation can help get you started.
	PCB 026 38444-81-4   DTXSID4074778 Searched by Approved Name.	
Details	Executive Summary	
Executive Summary	Quantitative Risk Assessment Values	PhysChem Parameters •
Executive Summary	No IRIS values	
Properties	o No PPRTV values	5.0
Env. Fate/Transport	№ No EPA RSL values	
	o No RfD calculated	logKow log(BCF) log(VP)
	No RfC calculated ♦ IVIVE POD not calculated	
	INIVE POD not calculated	
	Quantitative Hazard Values	
ADME > IVIVE	oral POD values	
Exposure	S No inhalation POD values	
Plana di dua	🐼 No In vitro activity data	No ovicting
Bioactivity	Cancer Information	No existing
Similar Compounds	o cancer slope factor	3
GenRA	🙆 No cancer unit risk values	(traditional) information what are
Genka	o cancer data	(traditional) informationwhat are
Related Substances	Genotoxicity Data:predicted to be non-genotoxic L <sup>2</sup>	
0	Reproductive Toxicology	alternative sources of information
Synonyms	So reproductive toxicity data available	
Literature		
Links	Chronic Toxicology	that can be used to address data
LINKS	No chronic toxicity data available	that can be used to address data
	Subchronic Toxicology	
	🙍 No subchronic toxicity data available	gaps?
	Developmental Toxicology	gaps:
	No developmental toxicity data available	
	Acute Toxicology	
	🙍 No acute toxicity data available	
	Subacute Toxicology	
	🐼 No subacute toxicity data available	



### Structural Activity Relationships (SARs) and Structural Alerts (SAs)

 A SAR (or SA) is a (qualitative) association between a chemical substructure and the potential of a chemical containing the substructure to exhibit a certain biological effect

e.g. carcinogenicity alerts reflected in the Supramolecule from Ashby & Tennant (1988) Mut Res 204: 17-115

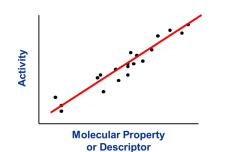


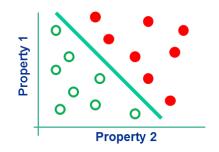


### Quantitative Structure-Activity Relationships (QSARs)

- A (Q)SAR attempts to relate (statistically or otherwise) the activity of one or more molecules to their physico-chemical properties or structural descriptors
- QSAR can be used to predict: Quantitative endpoints
   e.g. potency

Qualitative endpoints e.g. active / inactive







### **Collections of (Q)SARs**

- An Expert System is a formalised system, usually computerised that enables an end-user to make rational predictions of toxicity based on structure alone
- Expert systems are typically categorised by whether they are underpinned by:
  - empirically based algorithms such as QSARs e.g., TEST, OPERA
  - knowledge bases such as SARs e.g., Derek Nexus, Toxtree
  - or a hybrid e.g., TIMES, ChemTunes



## **Regulatory Applications of (Q)SARs**

- "Packaged mature knowledge for systematic reuse"
- For data gap filling to provide an estimate for a given (eco)toxicity/efate/phys chem endpoint in lieu of testing (replacement or supporting information)
- To rationalise spurious results in experimental data since the (Q)SAR is based on a larger body of data, provides a more compelling Weight of Evidence (WoE) to rationalise the validity of a potential outlier
- Essential for category development and associated read-across justification to provide a context of endpoint mechanistic similarity
- To add another line of evidence as part of a WoE within the context of an Integrated Approaches to Testing & Assessment (IATA)



### Scientific Validity: OECD Principles for (Q)SAR Validation

- A (Q)SAR should be associated with the following information:
  - a defined endpoint
  - an unambiguous algorithm
  - a defined applicability domain
  - appropriate measures of goodness-of-fit, robustness and predictivity
  - a mechanistic interpretation, if possible
- Principles were agreed by OECD in 2004 and associated guidance was published in 2007

Many QSARs/Expert systems use these principles as a basis to demonstrate potential utility for application. Reporting Formats (QMRF and QPRFs) exist to help summarise model characteristics and substance specific predictions.



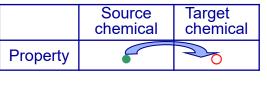
# QSARs which provide physchem (property) information

npTox Chemica	als Dashboard Home Sea	ch - Lists - About - Tools -				Submit Comments	Search a
		The r	ew Dashboard is a complete rebuild and is r	ew EPA CompTox Chemicals Dashboard eplacing the CompTox Chemicals Dashboard releas nentation can help get you started.	ed on July 12th 2020.		
	384	<b>3 026</b> 44-81-4   DTXSID407477 hed by Approved Name.	3				
s	Properties: LogKow: Octa	nol-Water					
tive Summary	LogKow: Octanol-Water	✓ Q Search	Chemical Properties				
rties							
ate/Transport	LEXPORT -			Summary			
	Туре	=   Average	≡   Median	≡   Range	≡ Unit	(#)	
	Experimental	5.76	5.76	5.76			
	Predicted	5.62	5.60	5.51 to 5.76			
> IVIVE	LEXPORT -			Experimental			
re	Source			Experimental Details		=	
vity 👗							
Compounds	PhysPropNCCT		5.76	J			
1	Yalkowsky et al. Chemosphere 2002, 48, 487-	-509	5.76				
l Substances				Predicted			
	LEUPORT *			Fledicied			
ms	Source	= Result	Calculation Details				=
ure 👗							
	ACD/Labs	5.51	Not Available	Not Available			
	ACD/Labs Consensus	5.52	Not Available	Not Available			
ents	EPISUITE	5.69	Not Available	Not Available			
	OPERA	5.76	OPERA Calculation Report [Inside AD]	Available			

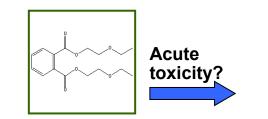


### **Read-across**

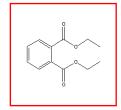
- <u>Read-across</u> describes the method of filling a data gap whereby a chemical with existing data values is used to make a prediction for a 'similar' chemical.
- Used within analogue and category approaches.
- A <u>target chemical</u> is a chemical which has a data gap that needs to be filled i.e. the subject of the read-across.
- A <u>source analogue</u> 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
- O Missing data



Known to be harmful



Predicted to be harmful



## **Ongoing issues with read-across**

- Although there is a wealth of technical guidance on how to develop read-across assessments, acceptance remains an issue. This is also not helped by the fact that read-across is typically an expert driven assessment.
- One issue impeding acceptance relates to the "uncertainty of the read-across prediction".
- 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, evaluating the utility of New Approach Methods (NAMs).
- Quantifying uncertainty and performance of read-across is a need as are approaches to more effectively characterise similarity contexts beyond structure e.g., metabolism, reactivity etc.

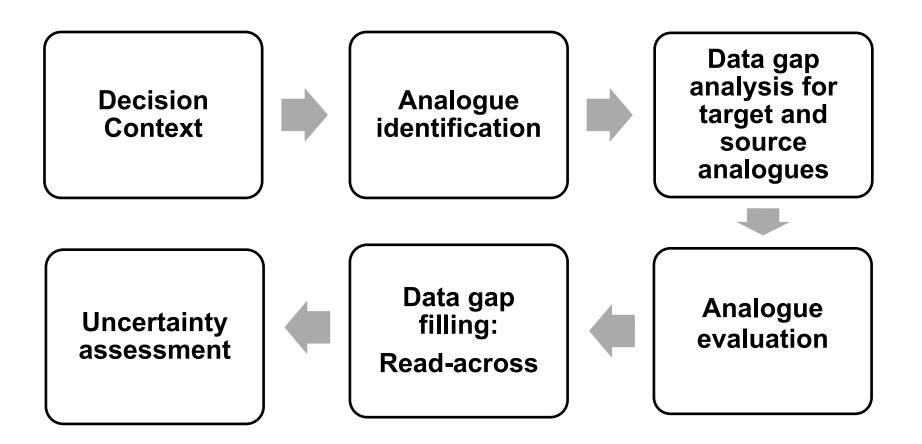


## **Read-Across Software Tools**

		Computational Toxicology 3 (2017) 1-18								
		Contents lists available at ScienceDirect								
5-5-2-2-		Computational Toxicology	Summary of key features of	selected publicly available read	-across tools.					
				AIM	ToxMatch	Ambit	OECD Toolbox	CBRA	ToxRead	CliPro
ELSEVIER	jou	rnal homepage: www.elsevier.com/locate/comtox	Development timeline	Java based version is dated 2012. Initial development of web version was 2005.	First public version released in Dec 2006	Original AMBIT tool was developed in 2004–2005	Proof of concept released in 2008	Implementation of the Low et al. [27]	Implementation of Gini et al. [22]	Implementation described in Russo et al. [45]
Navigating thr	rough the 1	minefield of read-across tools: A review of in	Type of Tool	Standalone	Standalone	Web-based and standalone	Standalone or Client/Server	Standalone	Standalone	Web-based
silico tools for	grouping	Q.	Latest Version	1.01 (Nov 2013) Static	1.07 (Jan 2009) Static	3.0.3 Ongoing	3.4 (July 2016) Version 4 released April 2017	0.75 First release	0.11 BETA Ongoing	First release
	,	man <sup>a,b</sup> , Prachi Pradeep <sup>a,b</sup> , Imran Shah <sup>a</sup>				Enhanced in 2013–2015	Ongoing			
<sup>a</sup> National Center for Computati 109 TW Alexander Dr, Research <sup>b</sup> Oak Ridge Institute for Science	ch Triangle Park (RTP),		Developed by	SRC Inc	Ideaconsult Ltd	ideaconsult Ltd	LMC, Bourgas	Fourches Lab at North Carolina State University	Istituto di Ricerche Farmacologiche Mario Negri	Zhu Research Group at Rutgers University
Article history: Read- Received 29 March 2017 regul	A B S T R A C T Read-across is a popular data gap filling technique used within analogue and category regulatory purposes. In recent years there have been many efforts focused on the chal	Available from	https://www.epa.gov/tsca- screening-tools/analog- identification-methodology- aim-tool	https://eurl-ecvam.jrc.ec.europa. eu/laboratories-research/ predictive_toxicology/ esar_tools/toxmatch	http://cefic-lri.org/ lri_toolbox/ambit/	www.qsartoolbox.org	http://www.lourches- laboratory.com/software	http:// www.toxread.eu/	http://ciipro. rutgers.edu/	
Accepted 25 May 2017 Available online 29 May 2017 Keywords:	-	in read-across development, its scientific justification and documentation. Tools have als oped to facilitate read-across development and application. Here, we describe a number of J able read-across tools in the context of the category/analogue workflow and review th capabilities, strengths and weaknesses. No single tool addresses all aspects of the workflow.		CAS, Name, SMILES, structure drawing/import	CAS, Name, SMILES, InChi	Name, identifiers, SMILES, InChl	CAS, Name, SMILES, structure drawing, MOL, sdf	Mol file, descriptors as txt	SMILES	PubChem CID, CAS, IUPAC, SMILES, InChI
Acyworus: Category approach Analogue approach Data gap filling Read-across		how the different tools complement each other and some of the opportunities for their fur ment to address the continued evolution of read-across. Published by	tatapoint coverage	N/A	Any based on user input	IUCLID <sup>2</sup> 5-supported endpoints (43 total)	Any as per the regulatory endpoints	Any based on user input	Mutagenicity and Bioconcentration Factor (BCF)	Any based on user input
(Q)SAR Trend analysis Nearest neighbours			Analogue Identification Approach	Fragment matching	Distance and correlation based similarity indices based on descriptors or fingerprints	Substructure or similarity searching using structure, name, SMILES, InChi	Category definition followed by subcategorisations	Tanimoto distance using chemical and biological descriptors	VEGA similarity algorithm	Weighted Estimated Biological Similarity
(Patlewicz	et al., 2	017)	Neighbour Selection	Automatic	Automatic	Manual	Automatic + Manual Filter	Automatic	Automatic	Automatic + Manual Filter
			Data Source	Tool provides inventory index	User provided or tool provided	User and tool provided	User provided or tool provided	User provided	Tool provided as a result of the EU ANTARES project	User provided but tool provides PubChem in vitro data
			Quantitative vs Qualitative	N/A	Both	User determined - Qualitative	Both	Qualitative	Qualitative for mutagenicity, quantitative for BCF	Qualitative
			Visualisation	None	Standard 2D plots, histograms and similarity matrix	None	Standard 2D Plots	Radial plot of neighbours	Interactive Neighbour plot	Activity Plot
			Output/Export	Output reports in the form of HTML, pdf or Excel	sdf or txt files of data, image files of plots	Assessment report as docx or xlsx, data matrix as xlsx	IUCLID format, pdf and rtf files of prediction report, text files of data, image files of plots etc	NA	Image file of plot	Tabulation of predictions and image of similarity plot
			<sup>2</sup> IUCUD stands for later	ational Uniform Chamical Infor	mation Database IUCUD is a softwar	re program for the administr	cation of data on chamical substan	car first developed to fulfill	El information mou	icomonte undor PEACU

<sup>2</sup> IUCLID stands for International Uniform Chemical Information Database. IUCLID is a software program for the administration of data on chemical substances first developed to fulfill EU information requirements under REACH.







## A Harmonised Hybrid Read-Across Workflow

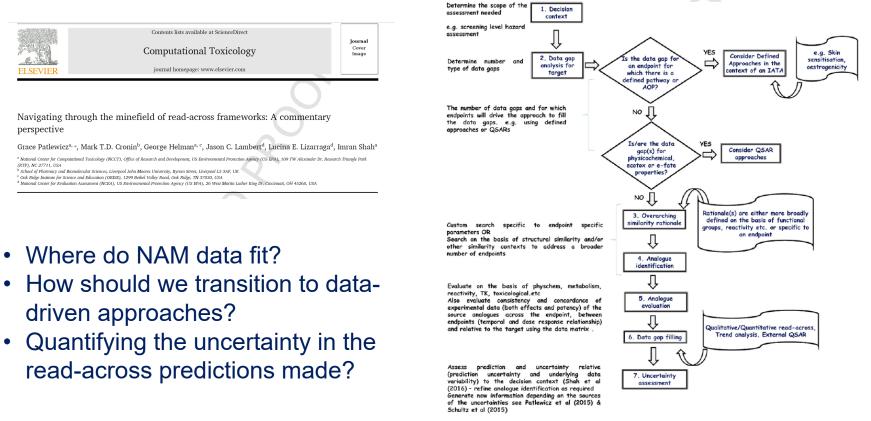


Fig. 9. A harmonised hybrid development and assessment framework.

Patlewicz et al., 2018



### **Data-driven read-across approaches**

ARTI

GenRA (Generalised Read-Across) Predicting toxicity as a similarity-weighted activity of nearest neighbours based on chemistry and bioactivity descriptors (Shah et al, 2016)

 Goal: To establish an objective performance baseline for read-across and quantify the uncertainty in the predictions made

	Regulatory Texicology and Pharmacology 79 (2016) 12-24	
ELSEVIER	Contents lists available at ScienceOfrect Regulatory Toxicology and Pharmacology journal homepage: www.elaevier.com/locale/yrtph	Regulatory Bischoolso and Planmac voluge Ware ware the
	evaluating read-across prediction and performance alidity approach characterized by chemical structure	Constituti



and bioactivity information

Imran Shah <sup>a,\*</sup>, lie Liu <sup>b, c</sup>, Richard S, Judson <sup>a</sup>, Russell S, Thomas <sup>a</sup>, Grace Patlewicz

icology, Office of Research and Development, U.S. Environmental P tion Science, University of Arkansas at Little Rock, AR 72204, USA

In Sympositic Control of Control of Control of Computational Toxicology, Office of Research and Development, U.S. Env Agency, Research Triangle Park, NC 27711, USA

С	L	E	I	N	F	0		۸	B	s	т	R	A	С	т	

Article history: Received 25 September 2015 Received in revised form 20 April 2016 Accepted 3 May 2016 Available online 9 May 2016	Bod-screen is a perpediar data, gap filling technique within category and analogue approaches for regu- latory pargueses. A necessaria of market and the screen scr
Krywords: Read-across Nearest mighbors Local validity domains (Q)SAR K/N Bioactivity TosCast	Over 3229 different chemical structure descriptors were generated for a set of 1778 chemicals and supplemented with the outcomes from REI in vite assays. The read-accous prediction of buckity for role chemicals with in vive data was based on the similarity weighted endpoint outcomes of its nonest regiptors. The approach endeld a proformance baseline for read-accous predictions of specific shugh outcomes to be established. Bioactivity descriptors were offices found to be more predictive of in vive trainity on the set of the structure of the set of the set of the set of the set of the set chemical proformation. The set of the set of the set of the set of the set of

$$y_{i}^{\beta,\alpha} = \frac{\sum_{j}^{k} S_{ij}^{\alpha} x_{j}^{\beta}}{\sum_{j}^{k} S_{ij}^{\alpha}}$$

Jaccard similarity:

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

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

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

 $v_i$  = predicted activity of chemical  $(c_i)$ 

 $x_i^{\beta} = activity of c_i in \beta$ 

 $s_{ii}^{\alpha} = Jacccard similarity between x_i^{\alpha}, x_i^{\alpha}$ 

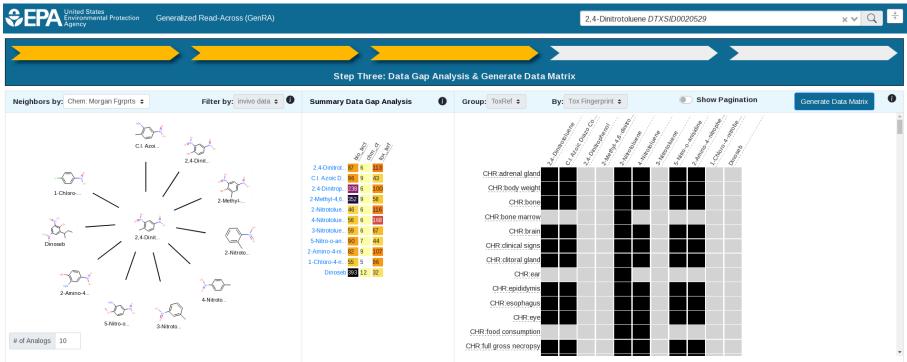
k = up to k nearest neighbours







# Search for a chemical of interest (target) using the search box



Radial plot with target in the centre and source analogues (similar) ordered clockwise by decreasing similarity (Jaccard)



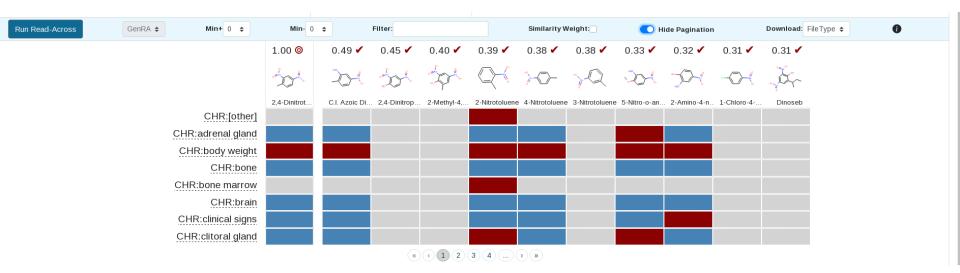


- How data poor is my target and what data exists for the source analogues identified
- Do they address the data gaps of interest for the target chemical?



What is the consistency and concordance across my source analogues? Should I deselect analogues from consideration from the entire set of predictions? Should I consider subcategorising the analogues selected?

Toxicity data represented as binary outcomes – red (positive), blue (negative), grey (no data)





Run Read-Across	Min+ 0 ≑	Min- 0	\$	Filter:			Similarity W	/eight:	💽 ні	ide Pagination		Download: File1	Type 🗢 🚺	
		1.00 🎯	0.49 🗸	0.45 🖌	0.40 🖌	0.39 🖌	0.38 🗸	0.38 🗸	0.33 🗸	0.32 🖌	0.31 🗸	0.31 🗸		
							<u>}-</u>	, Q				- d-		
		2,4-Dinitrot	C.I. Azoic Di	. 2,4-Dinitrop	2-Methyl-4,	2-Nitrotoluene	4-Nitrotoluene	3-Nitrotoluene	9 5-Nitro-o-an	2-Amino-4-n	1-Chloro-4	Dinoseb		
	CHR:[other]													
	CHR:adrenal gland													
	CHR:body weight													
	CHR:bone CHR:bone marrow													
	CHR:brain													
	CHR:clinical signs													
	CHR:clitoral gland													
				Filter:			Circulturation					Devente est. Eile:	Туре ≑ 🚺	
Run Read-Across GenRA 🖨	Min+ 0 ¢	Min-					Similarity		_	lide Paginatio		Download: File	Type •	
First column is	Min+ 0 ¢	1.00 <b>(</b>	0.49 🗸	0.45 <b>✓</b>	0.40 🗸	0.39 🗸	0.38 <b>✓</b>	0.38 <b>✓</b>	0.33 🗸	0.32 🗸	0.31 🗸	0.31 🗸		
First column is	Min+ 0 ¢	1.00 ©			0.40 🗸	De	0.38 🗸	0.38 🗸	0.33 🗸	0.32 🗸	0.31 🗸	0.31 🗸	The t	
First column is updated with	Min+ 0 ¢	1.00 ©	0.49 🗸	0.45 🗸			0.38 🗸	0.38 🗸	0.33 ✔ ->>+	0.32 ✔ *	0.31 🗸	0.31 ✔	in the second seco	
First column is		1.00 ©	0.49 🗸				0.38 🗸	0.38 🗸	0.33 🗸	0.32 ✔ *	0.31 🗸	0.31 ✔		
First column is updated with	CHR:[other]	1.00 ©	0.49 🗸	0.45 🗸			0.38 🗸	0.38 🗸	0.33 ✔ ->>+	0.32 ✔ *	0.31 🗸	0.31 ✔	in the second	
First column is updated with	CHR:[other]	1.00 <b>◎</b>	0.49 🗸	0.45 🗸			0.38 🗸	0.38 🗸	0.33 ✔ ->>+	0.32 ✔ *	0.31 🗸	0.31 ✔		
First column is updated with	CHR:[other] CHR:adrenal gland	1.00 <b>◎</b>	0.49 🗸	0.45 🗸			0.38 🗸	0.38 🗸	0.33 ✔ ->>+	0.32 ✔ *	0.31 🗸	0.31 ✔		
First column is updated with	CHR:[other] CHR:adrenal gland CHR:body weight	1.00 <b>◎</b>	0.49 🗸	0.45 🗸			0.38 🗸	0.38 🗸	0.33 ✔ ->>+	0.32 ✔ *	0.31 🗸	0.31 ✔		
First column is updated with	CHR:[other] CHR:adrenal gland CHR:body weight CHR:bone	1.00 <b>◎</b>	0.49 🗸	0.45 🗸			0.38 🗸	0.38 🗸	0.33 ✔ ->>+	0.32 ✔ *	0.31 🗸	0.31 ✔		
First column is updated with	CHR:[other] CHR:adrenal gland CHR:body weight CHR:bone CHR:bone marrow	1.00 <b>◎</b>	0.49 🗸	0.45 🗸			0.38 🗸	0.38 🗸	0.33 ✔ ->>+	0.32 ✔ *	0.31 🗸	0.31 ✔		
First column is updated with	CHR:[other] CHR:adrenal gland CHR:body weight CHR:bone CHR:bone marrow CHR:brain	1.00 <b>◎</b>	0.49 🗸	0.45 🗸			0.38 🗸	0.38 🗸	0.33 ✔ ->>+	0.32 ✔ *	0.31 🗸	0.31 ✔		



## **GenRA – Current research**

- Consideration of other information to define and refine the analogue selection & evaluation
  - physicochemical similarity (Helman et al, 2018)
  - metabolic similarity (Boyce et al, 2022),
  - reactivity similarity (Nelms et al 2018)
  - transcriptomics similarity (Tate et al, 2021)
- Transitioning to quantitative predictions of toxicity
  - Using GenRA to predict Lowest Observed Adverse Effect Level (LOAEL), acute oral (median lethal dose) LD50 (Helman et al 2019a,b)
- Developing a compendium of expert driven read-across examples to investigate how data driven read-across with NAM data can mirror expert assessments (*in prep*)



## **Relative Potency Values**

- Relative potency values have been applied in the assessment of mixtures as described in the first talk of this session. They represent a special type of grouping approach as described in the existing OECD grouping guidance.
- Well known examples include Toxic Equivalent Factors (TEFs) which have been used to assess mixtures of dioxins and furans.
- However, there are specific caveats and criteria for the use of these TEFs.
- TEFs and the estimation of toxic units for mixtures of chemicals which contribute to a biological effect through a common toxicity pathway.



## **Relative Potency Values**

- In the TEQ approach, the most toxicologically relevant compound is used as the reference compound. Components of the mixture should act by the same single toxic pathway and be of the same compound type (structural/functional group similarity) as the reference.
- The components of the mixture are each assigned TEFs such that their individual toxicity is expressed as a fraction of the toxicity of the reference which is given a TEF of 1.



• TEF (component A) = <u>Reference effect value</u>

Component A effect value

- An example of an effect value (or "effect level") would be a LOAEL
- The amount of each component in the mixture is then multiplied by its respective TEF and the values for each component are summed to give an overall toxic equivalency relative to the reference compound
- TEQ = sum(concentration X TEF)
- But what if the effect value of Component A is missing?
  - This is where read-across, QSARs can play a role in filling in the missing gaps.

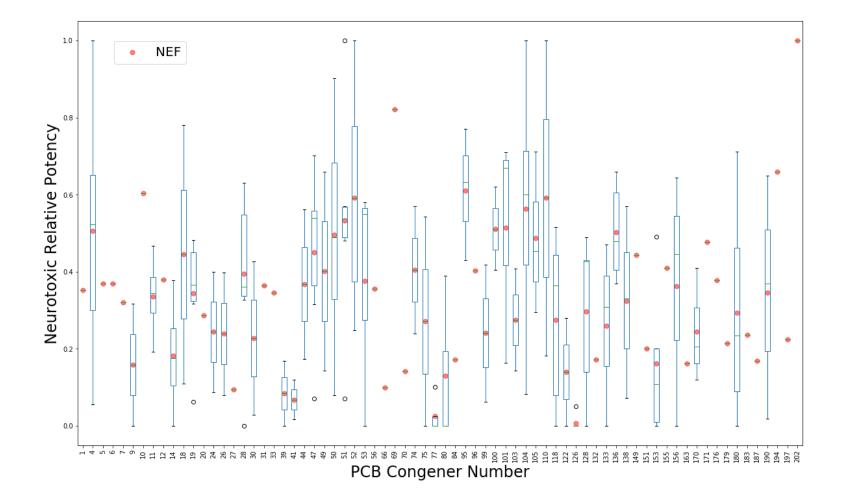




- Same principle as TEFs but NEFs were derived for the neurotoxicity of PCBs
- First developed by Simon et al (2007) who developed neurotoxic equivalent values for a dataset of 87 PCB congeners of which 83 congeners had *in vitro* experimental data
- However, the data was taken from several different studies each of which measured different effects. A more flexible interpretation of the TEF approach. Subsets of the 83 PCB congeners did overlap in terms of their *in vitro* data.
- Pradeep et al (2018) sought 1) to re-evaluate an alternative NEF from an expanded dataset and 2) investigate the feasibility of developing new QSAR models to predict NEFs.
- The resulting model could then be applied to estimate NEFs of the remaining untested PCB congeners



## Variability of NEFs



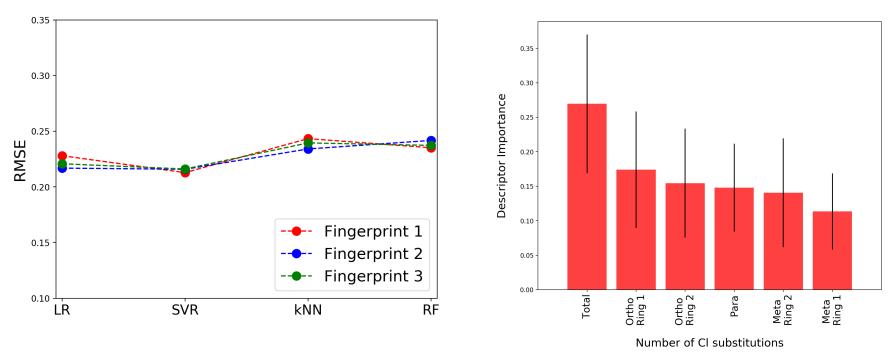


# **A QSAR model to predict NEFs**

- Is there a mathematical relationship between chemical characteristics and NEFs ?
  - Step 1: Characterise the PCB congeners in terms of structural characteristics using custom fingerprints (encode the chlorine substitution pattern of PCBs as a bitstring of 1s and Os)
  - Step 2: Investigate the feasibility of using different approaches to build QSAR models that relate the calculated inputs from Step 1 to known NEFs.
  - Step 3: Evaluate the robustness and performance of any QSAR models
- QSAR models derived had low predictivity (RMSE ~0.24) which was largely attributed to the large uncertainties of the data and the associated NEF values.
- Nonetheless, in the absence of better information, the derived NEFs and the QSAR predicted NEFs could be helpful to fill data gaps if applied with caution.



## **A QSAR model to predict NEFs**



4 different modelling approaches attempted

Which structural features were most influential in estimating the NEFs



- Computational toxicology covers a broad spectrum of different approaches
- Have highlighted a few of the main approaches to provide context
- The EPA CompTox Chemicals Dashboard provides a wealth of information (predicted and experimental data for hazard and exposure) which is a relevant starting point in the assessment of any substance of interest.
- Relative potency values are a special case of grouping approaches (read-across).
- Illustrated one case study where a QSAR model was developed to predict relative potency values for neurotoxicity (extending the socalled NEFs that Simon et al established) using chemical structural characteristics which could be applied to estimate NEFs for untested PCB congeners.