DATA INTEGRATION AND VISUALIZATION FOR TRANSPARENT COMMUNICATION IN READ-ACROSS: A GLYCOL ETHER CASE STUDY

A Thesis

by

MELINDA R. WILSON

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Chair of Committee,	Ivan I. Rusyn
Committee Members,	Weihsueh A. Chiu
	Natalie M. Johnson
Interdisciplinary Faculty Chair,	Ivan I. Rusyn

August 2016

Major Subject: Toxicology

ABSTRACT

Integration and communication of diverse information into human health assessments is often challenging given the wide variety of data types and sources under consideration. In this study, several new approaches are utilized for compiling and visualizing chemical structure and toxicological assay-derived information in support of both category and analogue read-across. A case study of propylene-based (P-series) glycol ethers was chosen, representing high production volume chemicals with a large database that includes physicochemical properties and various measures of mammalian toxicity. As a comparator, ethylene-based (E-series) glycol ethers were used to examine the potential for category separation and broader structural groupings. Two approaches for integrating and visualizing these complex data were used, Toxicological Priority Index (ToxPi) and Chemical-Biological Read-Across (CBRA), which allow for the unique incorporation of various types of data, differential weighting, relative ranking, and step retention to maximize transparency. The raw toxicity study data were transformed, scaled, and rendered to compose visualizations with alternate category grouping scenarios. We found overall that the glycol ethers group together within their structure-based category, as well as within series. Both ToxPi and CBRA facilitate effective communication of complex data and enable groupings and selection of most suitable analogues for read-across.

TABLE OF CONTENTS

Page

LIST OF FIGURES

Page

LIST OF TABLES

Page

1. INTRODUCTION

Whether to reduce the reliance on *in vivo* test models, build a constructive case for the application of read-across, or to develop a compliant submission for regulatory approval, the demand for integration and communication of modeled and experimentally-derived data has become a common challenge. In addition, most commodity chemicals that are being registered under the EU REACH (Registration, Evaluation, Authorization and Restriction of Chemicals) Regulation have some data gaps, primarily the information from experimental toxicity studies in animal models. These gaps need to be addressed before a robust safety assessment can be made. Read-across has become the go-to method for filling information gaps for one substance based on the data of another substance; usually one that is structurally similar, contains the same functional group, or exhibits similar toxicological and/or environmental fate properties ((Berggren et al., 2015), European Chemicals Agency, 2011). Often when implementing read-across, data from structure activity relationships (SARs) are utilized as predictors based on an analog's behavior under a given set of conditions, molecular descriptors, known activity, and prior testing of the analog(s) (Niska et al., 2008).

The European Chemicals Agency (ECHA) points out repeatedly that "structural similarity is a necessary [but not sufficient] pre-requisite for any prediction based on read-across under REACH" (European Chemicals Agency, 2015). Furthermore, it is advised in the same document that "for a read-across adaptation to be assessed and potentially accepted by the Agency, registrants have to show with clear reasoning and supporting data, that the substances involved in the read-across are structurally similar and are likely to have similar properties (or follow a similar pattern)." Thus, among many challenges in building a convincing read-across case, two are of a particular prominence to both the regulated

community and the regulators: (i) the inherent limitations of relying solely on structure and physico-chemical property information to justify "similarity", and (ii) the diversity of the information comprising a read-across application given the wide variety of data types and sources used to justify the "similar properties (or follow a similar pattern)" clause. Though each successive submission strives to improve the read-across application method, there are hindrances to creating a full and accurate representation from modeled data alone (Ball et al., 2014, Niska et al., 2008, Patlewicz et al., 2014). It has been suggested that some of these drawbacks can be mediated by the addition of *in vitro* to *in vivo* data (Low et al., 2014). Others, such as Hewitt et al. in 2010, have pointed out the benefits of integrating a combination of *in silico* models with a weight-of-evidence approach. Their use of (Q)SAR models such as CAESAR alongside the Derek expert system showed efficacy relevance, but echoed the limitations associated with an incomplete or limited dataset. (Hewitt et al., 2010) (Hewitt, Ellison, Enoch, Madden and Cronin, 2010) (Hewitt, Ellison, Enoch, Madden and Cronin, 2010)

Regardless of the data types that are used in read-across, it "must be, in all cases, justified scientifically and documented thoroughly" and "there may be several lines of evidence used to justify the read-across, with the aim of strengthening the case" (European Chemicals Agency, 2013). The requirement for "thorough documentation" of the scientific arguments that support the weight-of-evidence argument in read-across, or any other type of human health assessments, usually results in a voluminous amount of text, tables and charts that is difficult to communicate. The challenge of communicating the details of the extensive database that leads to a regulatory decision has been considered by several recent committees of the National Research Council (National Research Council, 2014, National Research

Council, 2015). This is increasingly pertinent as the 1976 Toxic Substances Control Act has undergone reform adding emphasis on utilizing emerging technologies to reduce and replace some of the traditional *in vivo* models. In addition, it calls for chemical grouping "into scientifically appropriate categories in cases in which testing of a chemical substance would provide scientifically valid and useful information on other chemical substances in the category." This synthesizes the increase in emerging data types with the need for integration leading to a read-across evaluation (H.R.2576, 2016). Several approaches for data integration have recently been proposed to handle challenges presented by high-dimensional toxicity data. Formal meta-analytic approaches are not immediately applicable because information from different data streams is not directly comparable. A variety of methods exist for weighting multiple streams of evidence differently, from largely qualitative, such as Hill criteria, to statistical frameworks and weighting schemes (Linkov et al., 2015). Examples of approaches to categorize chemicals that are both quantitative and incorporate expert judgment include the Toxicological Prioritization Index (ToxPi) (Reif et al., 2010, Reif et al., 2013) and Chemical-Biological Read-Across (CBRA) (Low et al., 2013).

In this study, we aimed to apply ToxPi and CBRA to grouping and analogue selection using a case study of glycol ethers. Glycol ethers are present in many products that we encounter daily, with their ubiquity ranging from pharmaceuticals and cosmetics to paints, adhesives and cleaning products (Oxygenated Solvent Producers Association (OSPA), 2012). They are synthesized by the reaction of an alkene oxide with a primary alcohol of varying lengths. The main chain can be lengthened by further reacting with excess alkene oxides to create mono-, di-, and tri- glycol ethers with methyl, ethyl, propyl and longer carbon groups. Not only is there similarity in structure, physicochemical properties, as well as environmental

and biological fate, but there are also extensive data on their toxicity from a variety of mammalian, *in vitro* and ecological models (European Centre for Ecotoxicology and Toxicology of Chemicals, 2005, The Dow Chemical Company, 2013). A previous study used read-across and tiered exposure assessment in risk assessment for one glycol ether, 1-methoxypropan-2-ol (PGME), as a representative for phase-in substances to be registered under REACH (Vink et al., 2010) and concluded that a robust data set provided by structurally related category members can predict missing values where needed. Here, we intend to expand these analyses to additional compounds in two broader categories, propylene (P-) and ethylene (E-) series glycol ethers and show how ToxPi and CBRA facilitate effective communication of complex data and enable groupings and selection of potential analogues suitable for a read-across application.

2. MATERIALS AND METHODS

2.1 Chemical Selection

The chemicals were chosen based on structural similarity, availability of data, and prevalence throughout industrial and consumer applications (see Table 1). Initially, eight P-series glycol ethers were chosen based on structural variety to better elucidate potential trends, and data availability to better represent and increase weight behind a trend, if found. After the initial integration, eight select E-series glycol ethers were added for comparison based on their structural similarity to the original P-series compounds.

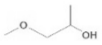
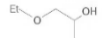
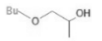
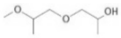
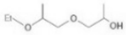
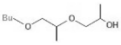
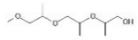
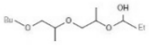
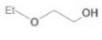
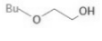
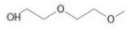
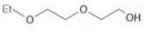
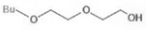
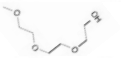
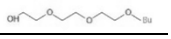
Chemical name	CAS No.	Abbrev.	Series	Oxide Units	-OH chain	Structure
1-methoxypropan-2-ol	107-98-2	PGME	P	Mono	methyl	
1-ethoxypropan-2-ol	1569-02-4	PGEE	P	Mono	ethyl	
1-butoxypropan-2-ol	5131-66-8	PGBE	P	Mono	butyl	
(2-methoxymethylethoxy)-1-propanol	34590-94-8	DPGME	P	Di	methyl	
(2-Ethoxymethylethoxy)propanol	30025-38-8	DPGEE	P	Di	ethyl	
1-(2-butoxy-1-methylethoxy)-propan-2-ol	29911-28-2	DPGBE	P	Di	butyl	
2-[2-(2-methoxy-1-methylethoxy)-1-methylethoxy] propan-1-ol	25498-49-1	TPGME	P	Tri	methyl	
1-[2-(2-butoxy-1-methylethoxy)-1-methylethoxy]propan-1-ol	55934-93-5	TPGBE	P	Tri	butyl	
1-hydroxy-2-methoxyethane	109-86-4	EGME	E	Mono	methyl	
2-ethoxyethanol	110-80-5	EGEE	E	Mono	ethyl	
2-butoxyethanol	111-76-2	EGBE	E	Mono	butyl	
2-(2-Methoxyethoxy)	111-77-3	DEGME	E	Di	methyl	
2-(2-Ethoxyethoxy)ethanol	111-90-0	DEGEE	E	Di	ethyl	
2-(2-Butoxyethoxy)ethanol	112-34-5	DEGBE	E	Di	butyl	
2-[2-(2-methoxyethoxy)ethoxy] ethanol	112-35-6	TEGME	E	Tri	methyl	
2-[2-(2-butoxyethoxy)ethoxy] ethanol	143-22-6	TEGBE	E	Tri	butyl	

Table 1. Glycol ether compounds analyzed in this study. A comprehensive collection organized by either propylene-based (P-) or ethylene-based (E-) series, oxide units (Mono-, Di-, Tri-), and then side chain (methyl, ethyl, butyl) with the structure shown on the right. Structures in the table are derived from the respective REACH dossier by CAS number. Where multiple constituents were available, the primary was chosen.

2.2 Data Acquisition and Transformation

First, the data from a variety of toxicological studies on selected compounds were collected from the disseminated REACH registration dossiers and assembled into a single spreadsheet matrix where each row was assigned to a specific chemical, and the individual columns represent distinct endpoints. Since many mammalian and ecotoxicity endpoints were evaluated for only a few compounds in the comparison, a criteria was set to only include parameters for which >40% of the data was available across the 16 compounds in the analysis, resulting in 21 endpoints being used for evaluation. This lead to certain traditional data being excluded, such as genotoxicity, to improve the competence of the overall comparison. The data for the additional E-series glycol ethers were also taken from the disseminated REACH dossiers documents (http://echa.europa.eu/web/guest/information-on-chemicals/registered-substances).

The data were compiled and organized into three test parameter categories: physicochemical properties, environmental toxicity, and mammalian toxicity. Again, due to data availability, only generated physicochemical and collected mammalian toxicity data were used. A modified Globally Harmonized System (GHS)-like classification for the categorization of chemicals was implemented (see Appendix Table 1) for endpoints which were not already given a published GHS hazard classification label. To equate the supplemented data to the existing GHS values, a reclassification system was devised to relate the raw data into a comparably scaled format. The resulting data were then normalized to be bounded by the interval [0, 1] to be compatible with the ToxPi input requirements. To ensure the transformations were an accurate representation of the original data, an original spreadsheet was maintained separate from the working spreadsheets and periodically checked

for accuracy to the original documents, as well as the working spreadsheets for the data representations. Each parameter is treated separately when rescaling, rendering it unitless, facilitating the integration of multiple data types by eliminating the need for unit conversions or other compensatory actions.

2.3 Toxicological Priority Index Data Integration and Visualization

The resulting toxicity data table was uploaded into ToxPi GUI ((Reif et al., 2013), available for download from http://comptox.us/toxpi.php) and slice composition determined by clustering of data types. The Mammalian Toxicity data rendered five categories as seen in Table 2 (with the number of representative endpoints in parenthesis): Acute Toxicity (5), Irritation (3), Reproductive Toxicity (4), Developmental Toxicity (3), and Repeat Dose Toxicity (6). The weighting scheme was chosen to reflect the number of parameters contributing to each slice; the more data inputs forming the slice, the larger the slice weight in composition, giving each endpoint an equal weight in the overall score. The ToxPi output not only consists of a pie chart representation for each chemical with the various endpoints as slices, but also includes a ranking of the chemicals, comparatively most to least toxic, based on the cumulative ToxPi score. After the information was processed, the outputs were organized in various ways to convey multiple aspects of the dataset and of the outcome. This was done by overlaying the resulting individual glycol ether ToxPis onto their corresponding data points on the graph of chemical rank vs. ToxPi score. This allowed for the combination of ToxPi output visuals to better elucidate potential trends. The resulting graphics were compared to the original dataset to ensure that the normalizations and visualizations had accurately rendered the information.

Toxicity Endpoint	P-Series Glycol Ethers								E-Series Glycol Ethers							
	PGME	PGEE	PGBE	DPGME	DPGEE	DPGBE	TPGME	TPGBE	EGME	EGEE	EGBE	DEGME	DEGEE	DEGBE	TEGME	TEGBE
Acute Toxicity																
LD_{50} (oral): Rat	0	0	0	0	0	0	0	0	.25	.25	.25	0	0	0	0	0
LC_{50} (inhalation): Rat	0	0	0	0	0	0	0	-	.25	.5	.25	.5	0	0	0	0
LD_{50} (dermal): Rabbit	0	0	0	0	0	0	0	0	.25	0	.25	0	0	0	0	0
Hematopoietic Toxicity	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
Hemolysis	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0
Irritation																
Skin Irritation	0	0	.75	0	0	0	0	0	0	.25	.75	0	0	.25	0	0
Eye Irritation	0	.75	.75	0	0	0	0	0	.25	.25	.75	0	.25	.75	0	1
Skin Sensitization	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-	-
Reproductive Toxicity																
NOAEL (oral): Rat	0	-	-	-	0	0	-	-	1	1	0	0	0	0	-	-
NOAEC (inhalation): Rodent	0	0	-	-	-	-	0	-	.75	.25	0	-	-	-	-	-
NOAEL (dermal): Rodent	-	-	-	-	-	.25	-	-	.25	-	-	-	-	-	-	-
Testicular Toxicity	0	0	0	0	0	0	0	0	1	1	0	1	1	0	0	0
Developmental Toxicity																
NOAEL (oral): Rat	0	-	-	-	-	0	-	-	1	.5	0	.75	0	0	0	0
NOAEC (inhalation): Rodent	0	0	-	0	0	-	0	-	1	1	0	-	-	-	-	-
NOAEL (dermal): Rodent	-	-	0	-	-	.25	-	-	.75	-	-	-	0	0	-	-
Repeat Dose Toxicity																
Sub-acute NOAEL (oral): Rat	0	0	0	.25	0	.25	-	.25	.75	.25	.5	.25	.25	.25	0	-
Sub-acute NOAEC (inhalation): Rat	0	0	0	0	-	.25	0	-	.5	0	0	-	.75	.5	-	-
Sub-acute NOAEL (dermal): Rabbit	0	-	-	0	-	-	-	-	.25	0	-	-	-	-	0	0
Sub-chronic NOAEL (oral): Rat	0	0	.25	-	.25	.25	-	.25	.75	.5	.75	-	.5	.5	0	-
Sub-chronic NOAEC (inhalation): Rat	0	0	0	0	-	-	0	-	.75	.25	.5	0	.75	.75	-	-
Sub-chronic NOAEL (dermal): Rabbit	0	0	.25	0	-	.25	.25	-	.25	-	.5	.75	.25	.5	0	-

Table 2. Rescaled mammalian toxicity data. A GHS-like format was applied to the values from the in vivo endpoints acquired (see Supplemental Table 1) to provide a binned evaluation, whose values were then rescaled within the range [0, 1].

2.4 Chemical Descriptors

For each of the glycol ethers, two methodologies were used to describe the compounds by quantitative structural attributes to provide a standardized assessment of inherent properties in common, such as bond strength or length. Constant and near constant descriptors as well as highly correlated descriptors were excluded and descriptor values were normalized on a scale from 0 to 1. The Chemistry Development Kit (CDK) produced 85

relevant chemical descriptors, while Dragon provided 535 for chemical comparison (Steinbeck et al., 2003; Mauri et al., 2006).

2.5 Principal Component Analysis

The chemical descriptor data were used to perform separate Principal Component Analyses to separate out the planes of greatest global separation and look for a structural correlation if a trend was found. The plots were created using the open source software environment R (R Core Team, 2015).

2.6 Chemical-Biological Read-Across Data Integration and Visualization

After formatting the biological descriptors generated from the mammalian toxicity data and chemical descriptors into separate text files, they were loaded into the CBRA graphical user interface (http://www.fourches-laboratory.com/#!software/vkttx) and visualized. The application's results show an individual chemical and the closest neighbors with respect to separate biological and chemical similarity. The visual created is a radial plot with the most similar chemicals closer to the center and top of each half of the chemical's representative plot. A calculated Tanimoto coefficient is the similarity determinant value and corresponding vector length. The outputs can be visually refined by number of neighbors or by Tanimoto coefficient threshold for each plot; however, we chose to present the data for the five nearest neighbors to the compound of interest.

3. RESULTS

The challenge surrounding confidence in the use of read-across analyses has been frequently scrutinized as lacking characterization of the factors that may contribute to uncertainty in the analogue (Patlewicz et al., 2013; Patlewicz et al., 2015). While most often the read-across analysis begins with identification of the chemical similarity (Wu et al., 2010), assessment of the comparative biological activity is usually a secondary consideration, if discussed at all. While the basis for read-across is a lack of information for a certain endpoint, data gaps exist for most compounds; the reality is a checkerboard of missing data even among the well-defined libraries of chemicals. Thus, we evaluated the overall "similarity" among two closely related groups of chemicals, P- and E-series glycol ethers (Table 1), in terms of their likeness when considering biological (i.e., mammalian toxicity studies), chemical (i.e., 2D chemical descriptors), and combined datasets.

3.1. Biological Similarity

We used a complement of mammalian toxicity data on the selected compounds (see Table 2) to rank, group, and produce integrated visualizations of the results.

First, we utilized the Toxicological Priority Index (ToxPi), a tool for transparent integration and visualization of data across disparate information domains (Reif et al., 2010; Reif et al., 2013). ToxPi produces a dimensionless index score that enables multiple sources of evidence to be integrated into visual rankings that are transparent and facilitate decision making. Specifically, data were translated into ToxPi scores for all compounds, first independently for P- or E-series glycol ethers (Figure 1), and then by combining all compounds into one analysis (Figure 2).

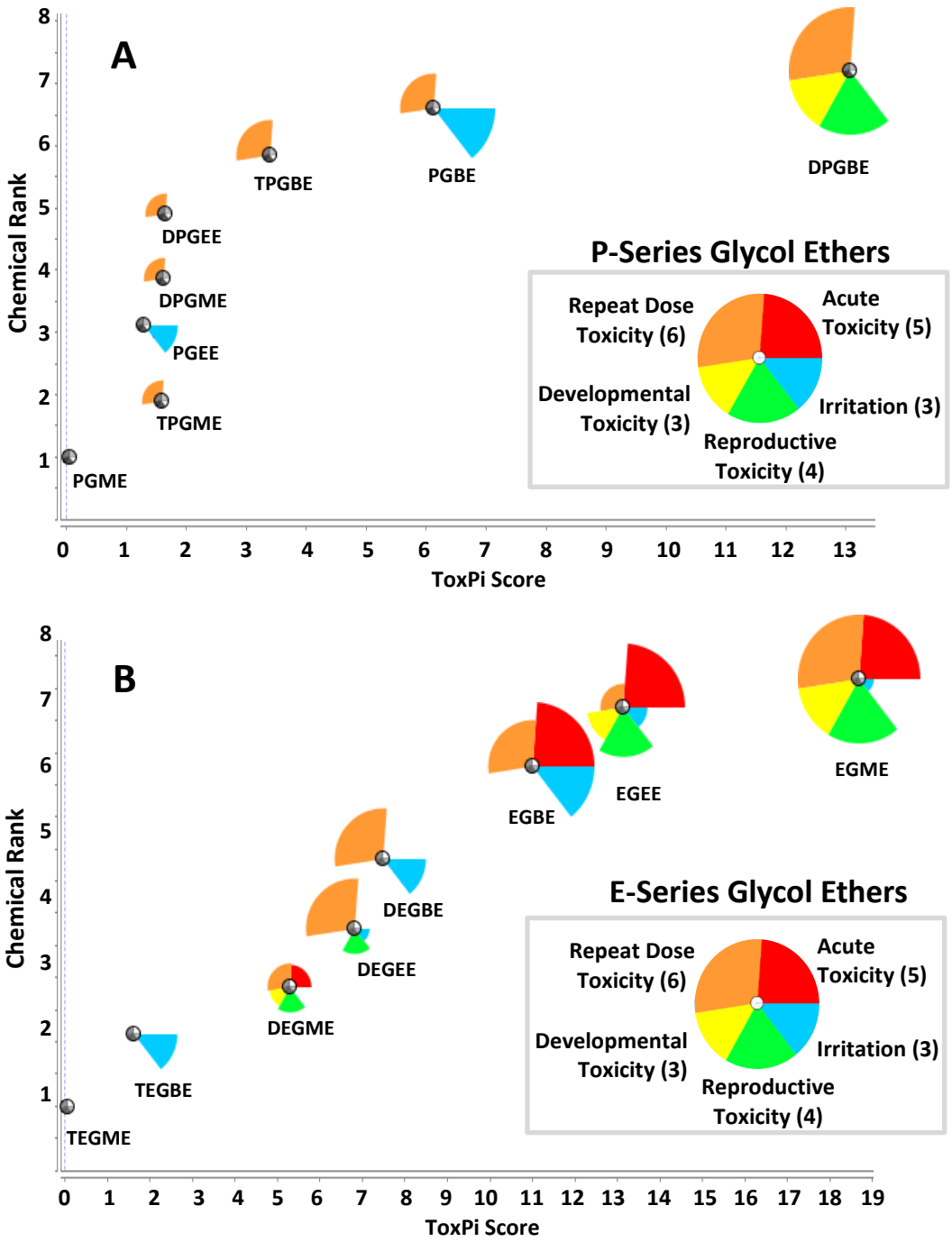


Figure 1. ToxPi analysis of separate E- and P-series glycol ethers. ToxPi scores were used to create a relative rank of chemicals within each series for categories of mammalian toxicity. Visualization of the ToxPi-based groupings and trends are seen within P-series glycol ethers (A) and E-series glycol ethers (B) score/rank graphs. The individual ToxPi image was overlaid onto the corresponding rank/score. The greatest determinant in the P-series comparison can be attributed to repeat dose toxicity. The largest similarity in E-series compounds comes from the intrinsic chemical properties (the quantity of groups, i.e., mono-, di- and tri-) and creates a spectrum rather than clusters. The inset shows the components of each slice: data type and number of endpoints (in parentheses).

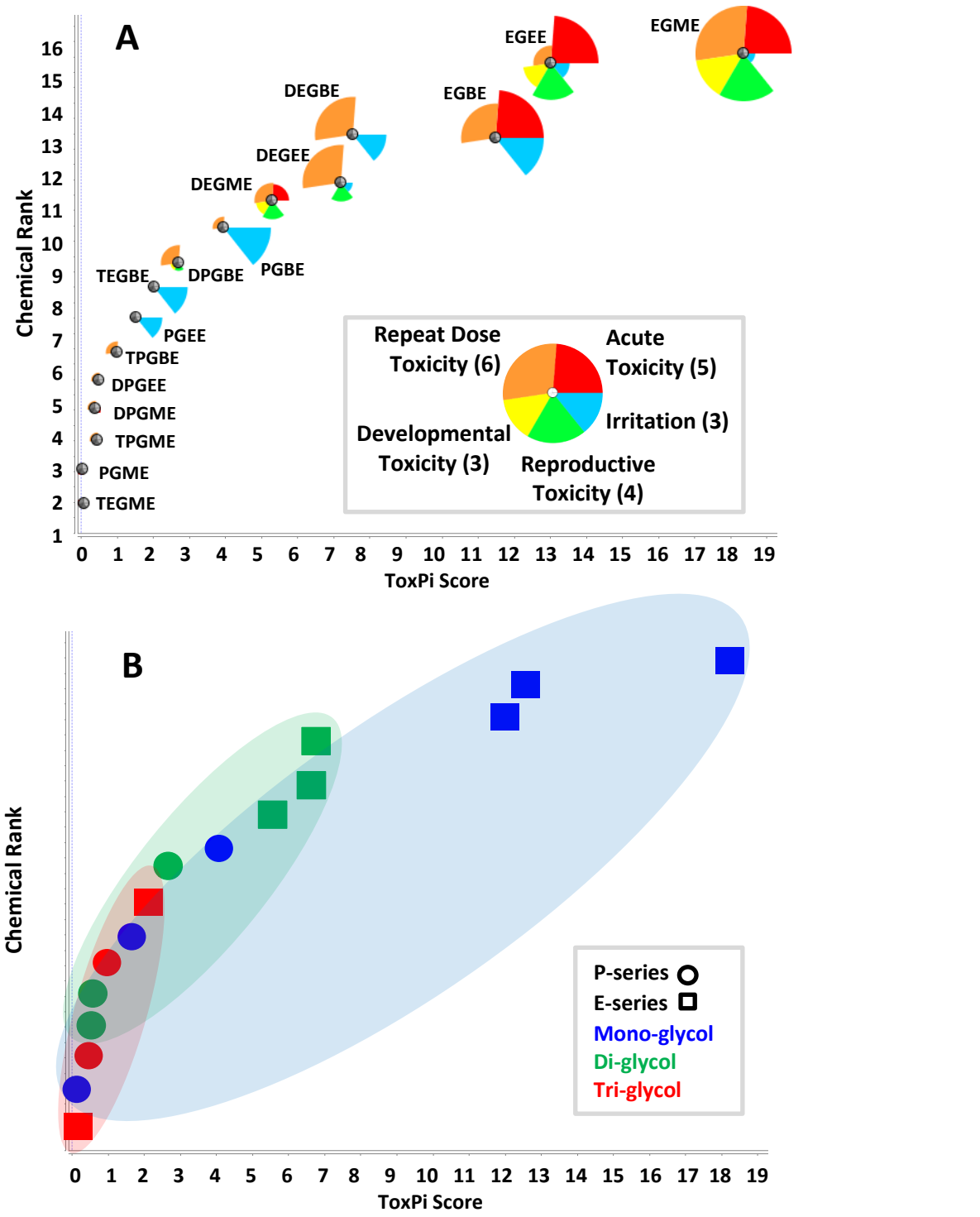


Figure 2. ToxPi analysis of combined P- and E-series glycol ethers. The ToxPi-based groupings among mono- (blue dots), di- (green dots), and tri- (red dots) glycol ethers are shown for the entire dataset, along with the individual ToxPi images (A) or clouds highlighting structural similarity (B). Even when combined with P-series, the E-series mono-glycols are still relatively more biologically active, especially within the Acute Toxicity (red slice) category. The demonstrated structure-activity relationships among glycol ether presents mono-glycols (blue cloud) as relatively most biologically active, with tri-glycols (red cloud) being the least, especially within their respective series (P- or E-).

The resulting ToxPi visualizations are comprised of both the pie chart representation of toxicity, as well as the graph comparing the chemical's rank (1 = least toxic, 8 or 16 = most toxic) to the ToxPi score (an integrated value of the attributed slice area). The larger the slice, the higher its biological activity in given mammalian toxicity tests comprising the slice. The small inner circle defined by the white-to-black gradient represents the percent missing data contributing to the ranking. The darker the shade of the inner slice, the more missing data, and therefore the less assurance that may be had in a true representation of toxicity rank.

First, by comparing the P-series and E-series glycol ethers separately within their own set, we observed several groupings and trends. In Figure 1A the butyl ethers (abbreviations ending in "BE") are relatively more toxic, largely due to high Acute and Repeat Dose Toxicity scores. Overall, the P-series cumulative toxicity profiles fall in a gradient rather than distinctly clustering by structural similarity. Conversely in Figure 1B, groupings based on wedge magnitude are apparent among the E-series ethers, illustrated by the red Acute Toxicity slice being relatively large only for the mono-glycols (EGME, EGEE, and EGBE). This shows the E-series separating by the mono-, di-, or tri-glycol chemical property, with the mono-glycols being overall more toxic, and the tri-glycols being the relatively least toxic.

Next, we combined the P- and E- series data and re-analyzed the ToxPi results using the same criteria and grouping of individual assays with the increased dataset. Two trends in the grouping of agents based on their biological activity are evident, largely in concert with the chemical similarity (Figures 2A-B). The P-series ToxPi labels are shown below their corresponding points while the E-series labels are displayed above the points. Figure 2A displays each individual chemical's ToxPi overlaying the corresponding Rank and Score.

Figure 2B layers the chemical similarity onto the toxicity data-based ToxPi ranking. The Acute Toxicity category is still driving the mono- E-series glycol ethers to be most toxic, even when combined with additional compounds. Figure 2B shows the E-series (squares) tended to have higher ToxPi scores (indicating relatively higher toxicity in the tests that were included in this analysis) than the P-series (circles) overall, which placed them further to the right on the x-axis. It is evident that there are sub-clusters of P- and E- series mono- (blue cloud), di- (green cloud) and tri-glycols (red cloud), with the mono-glycols exhibiting relatively higher toxicity for both sets overall.

3.2. Chemical Similarity

Because glycol ethers are relatively simple chemicals in terms of structure, we next sought to visualize the relationships among them using a high-dimensional dataset of 2D chemical structure descriptors that are commonly used to develop quantitative structure activity models. Rather than isolating specific functional groups, we used molecular descriptors that are a mathematical interpretation from the transformation of chemical information into an encoded representation of a molecule. This was done using the SMILES code for the major isomers, where applicable. Each compound is described by quantitative structural attributes developed for the purpose of providing a standardized description of structural properties that are common across chemicals and can be used to model structure-based commonalities and provide a chemical basis for potential toxicity relationships.

Figure 3 depicts a 2D approximation of the relative distance between all 16 compounds in a 85- (CDK descriptors) or 535- (Dragon descriptors) dimensional space using principal components analysis. The CDK descriptors (Figure 3A) show more than 84% of the

variation can be accounted for by the first two principal components, separating spatially first by glycol unit quantity, and then by series (P- vs. E-). Very similar trends were observed when using a different set of molecular descriptors. Dragon descriptor-derived visualization (Figure 3B) shows that most similarity among compounds is determined by the number of glycol units (mono-, di-, tri-) in the first principal component (explaining over 67% of the variability), and then by the series (P- vs. E-) in the second principal component (explaining an additional 14%). Combined, the first two components describe much of the variation (>81%) in the data.

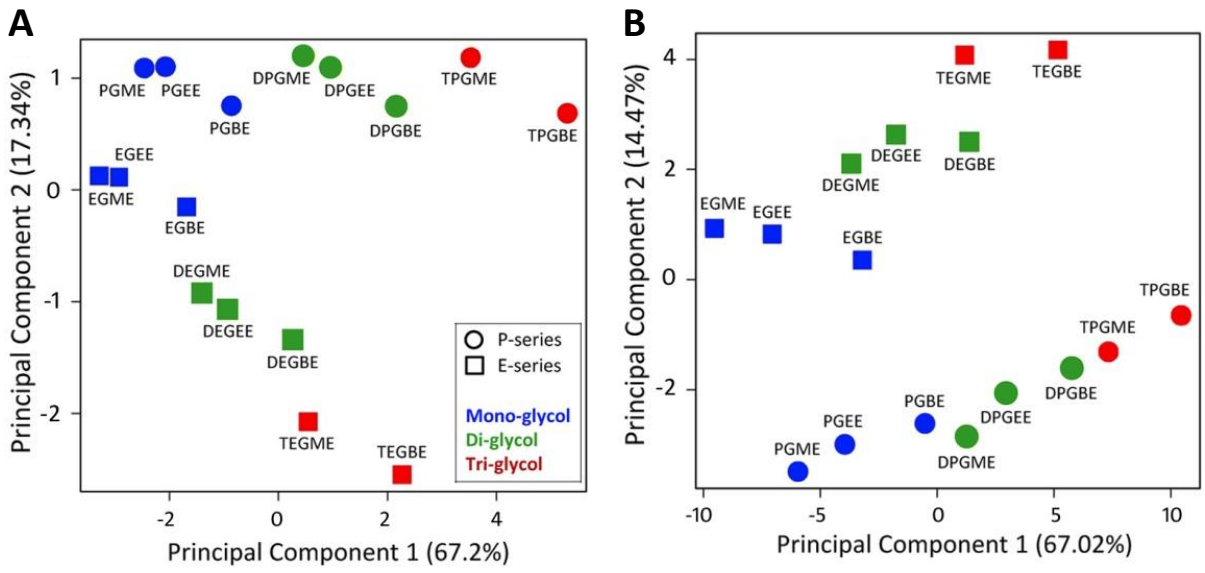


Figure 3. Principal component analyses for combined glycol ethers. The Chemistry Development Kit (A) provided descriptors which describe over 84% of the variance within the first two components. The DRAGON PCA (B) accounts for over 81% within the first two components. Chemical groupings are described in the inset legend by the glycol ether series (P: circle; E: square), as well as by structure: mono- (blue), di-(green), and tri-(red) glycols.

3.3. Combined Biological and Chemical Similarity

While ToxPi and PCA demonstrated concordance between biological and chemical-based similarity, they were most useful for identifying the groupings (or categories) within the selected compounds and classes. These are less useful in selecting single analogue(s) for read-across. Thus, we used another type of integrative analysis, an approach of chemical-biological read-across (Low et al., 2013). CBRA infers a compound's activity from those of its chemical and biological analogs. By directly relying on the experimentally-obtained biological measurements, as well as inherent physicochemical properties, the result is a computed similarity score based on other agents in the dataset and produces a radial plot showing the relative distance to the closest chemical "neighbors."

As opposed to groupings in ToxPi and PCA, CBRA allows for identification of most suitable "source" compounds (outer compounds) for any given "target" (the center of the radial plot). In the CBRA plot outputs (Figure 4), the left side represents the biologically similar compounds, while the right side shows the nearest neighbors with respect to chemical similarity, both calculated using the Tanimoto coefficient. The closer a chemical is to the center of the plot (i.e., to the "target" compound), the higher the Tanimoto coefficient, the more similar the two compounds are. The chemical abbreviation's accompanying shape and color corresponds to its known category assignment for the two parameters: glycol ether series type and chemical structure/composition (color). CBRA plots show that the nearest neighbors in chemical space tend to follow chemical composition (i.e., the number of glycol units, mono-, di-, tri-) rather than series, while the biological similarity echoes the ToxPi-derived rankings. The closest chemical neighbors also reflect the spatial distribution from the Dragon PCA plot (Figure 3B) from which the chemical component was sourced.

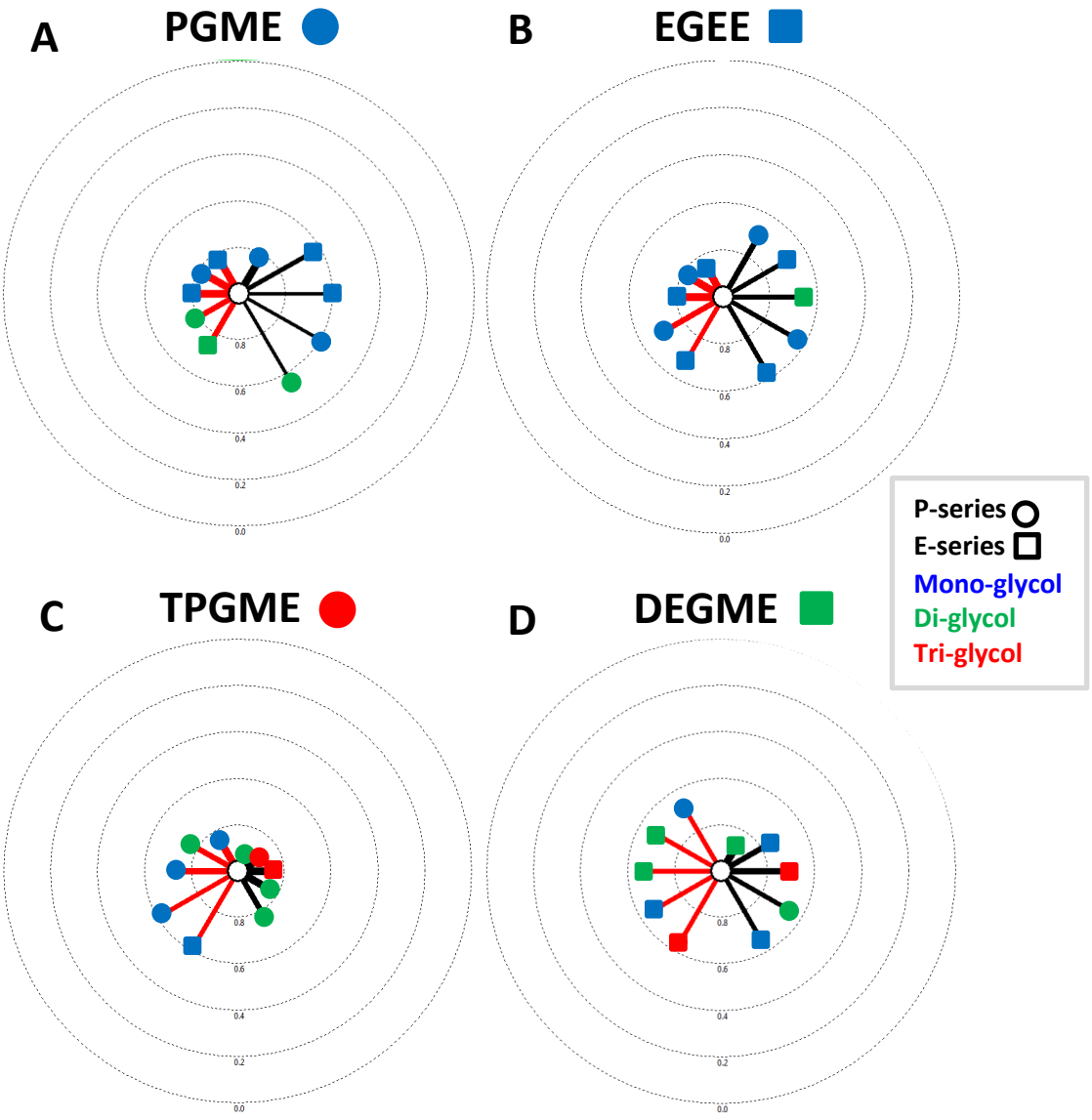


Figure 4. Selected CBRA radial plots. Figures A-D represent an individual chemical and their biologically closest (left) and chemically closest (right) neighbors. The color and shape accompanying the chemical abbreviation represent series (P: circle; E: square), as well as structure: mono- (blue), di-(green), and tri-(red) glycols, as presented in the inset legend. This also was applied to identify the chemical neighbors radiating from the plot's center. The distance is determined by the Tanimoto coefficient, therefore the shorter the distance to the center, the larger the coefficient, the more similar the compound is to the one of interest. Figure 4A and 4B represent a relatively similar prediction for the nearest neighbors, both identifying acceptable candidates also of the mono- class and of the same series. PGME (4A) shows the nearest biological neighbor being an E-series chemical, while EGEE (4B) follows the expected biological profile its most similar neighbor being an E-series mono-glycol. Figure 4C shows the radial plot for TPGME with the nearest neighbors in biological and chemical space as predominantly other P-series compounds. The nearest neighbor for DEGME (4D) in chemical space is another E-series di-glycol, which is also present as a close neighbor in biological space.

For example, Figure 4B shows that for a relatively more toxic compound within the dataset, EGEE, other E-series mono-glycols (blue squares) are included in the nearest neighbors with respect to both biological (left side) and chemical (right side) space. This echoes what is seen in the separate biological and chemical analyses, but is presented in a comprehensive figure. Figure 4C shows TPGME, a relatively nontoxic compound. Nearly all of the nearest neighbors in biological and chemical space are other P-series compounds, reaffirming that the P-series group as relatively less toxic than the E-series. The analysis for DEGME (Figure 4D) does provide potential for a chemical similarity argument, however the biological nearest neighbor is a P-series mono-glycol. Though the closest biological neighbor PGBE does have a relatively high Tanimoto coefficient score, it is still important to note the toxicity profile associated with the two compounds. While DEGME has low biological activity across a variety of toxicity categories, the PGBE profile is dominated by the Irritation category, something that DEGME is lacking. This further emphasizes the importance of a holistic and judicious evaluation of both chemical and biological similarity.

4. CONCLUSIONS

Programs such as ToxPi and CBRA have clear utility for category elucidation and formation, as well as comparative analyses within and across a product category. There are caveats and data considerations involved in constructively making use of these applications. A driving factor for any comparison is having a complete, or near-complete, dataset for the subjects of interest. Therefore, a hindrance to an informative output may be attributed to the diversity and inclusivity of data or lack thereof. Variations between study designs and endpoints observed can lead to a mismatched or incompatible dataset, which may produce a less informative representation as well as echo a need for broad and concurrent method standardization. When using a tool such as ToxPi, gaps in a dataset may pose an area for debate. Many compounds that have been screened resulted in no adverse effect observed, meaning they were pushed to the highest concentration or dose of the study, whether *in vitro* or *in vivo*, and had no negative effect observed. While this may be good news for a manufacturer, it leaves a large comparison study at a loss. The question arises on how to use the known data in the most informative, yet accurate, representation. Since no effect was seen, some may argue the endpoint value be purposefully left blank as not to attribute non-existent toxicity at the highest dose tested. A case can also be made for adding in the no observable adverse effect level (NOAEL) to complete the table, or the maximum tested dose to represent that there is no observed adverse effect at least up to that dose point. Since one can't be sure how high the dose could continue to produce no effect without a proper dose-response or point of departure analysis, expert judgement and documentation of the methods taken becomes a necessity. Leaving the data point blank decreases the predictive capabilities of the generated output since the lack of response is now in the same undefined nebulous

alongside the untested and potentially potent compounds. This leads to restrictions based on the individual study design and is apparent when combining various sources ranging across design and time.

CBRA application handles a lack of information in a slightly different way. Since the comparison is done with a Tanimoto coefficient comparing similarity and doesn't include a ranking, a missing data point doesn't affect an overall score, simply a less confident prediction in the most similar compounds to the one of interest. Another invitation for open communication between stakeholders arises with multiple data entries available for a certain chemical and endpoint. Should one use the most recent study, the study with the largest sample size, or the most conservative reported value? It is also important to note that compounds may have a similar ToxPi score, yet very different toxicity profiles. The holistic approach presented by ToxPi builds not only individual profiles, but identifies a bigger picture which may point to otherwise-overlooked trends in biological activity. The profile, score, and ranking all serve to emphasize the benefits of a multi-visual output. There are advantages and disadvantages to weighting options. In this study we chose to weight endpoints equally, with slice weight reflecting the number of endpoints. By assigning equal weight to all parameters, the acute toxicity as well as irritation data have the same impact on the overall toxicity profile. Depending on the priorities of stakeholders, preferential weighting may be assigned to a specific slice or category in ToxPi.

This glycol ether case study provided a relatively data-rich platform to work from, however even a data-limited compound may be of utility. When compared to a set for which there is a subjectively sufficient amount of associated data, the data-limited substance can be added, analyzed with ToxPi, and the resultant profile may help identify what category or

toxicity profile grouping it would best fall into. This would be an applicable method for a read-across type approach. Given a wide variety of known relatively toxic and non-toxic compounds, such as utilizing ToxCast data, one could incorporate a wide variety of parameters to build and visualize unique profiles of toxicity. When compared to a diverse and stratified set, even a partial match for a data-limited chemical may be useful in gauging potential bioactivity, groupings, trends, overall characterization, or direct future study. Similarly, such data could also be supplemented into CBRA, which in addition, would make use of its physicochemical properties to assign similarity to others within that set, giving another dimension for read-across comparison to establish sufficient comparability. These applications, particularly CBRA, can also be of use for emerging types of data, such as –omics results or "big data." A multitude of endpoints from *in vivo, in vitro,* and *in silico* can be entered into the biological CBRA component to aid in a read-across type approach for categorization based on combined chemical and biological similarity for a registration application. These tools are not intended to replace traditional methods, but rather an addition to strengthen a weight-of-evidence approach, increasing the confidence towards producing a quality assessment.

By allowing diverse data types to be used in conjunction, a global similarity can be evaluated, alongside the traditional endpoint-specific similarity. This analysis aimed to utilize ToxPi for clustering similarly-active compounds and identifying potential chemical drivers for the trend. Though some compounds showed similarity (for example, many butyl ethers scoring highly for Irritation Toxicity), the results highlight the importance of using less traditional data in conjunction to chemical description where available. This case study used very structurally comparable compounds, both within and across the glycol ether series'. The

wide spread of biological responses for standard endpoints emphasizes the need to look past simple chemical structure and incorporate biological activity profiles as much as possible in decision making. This urges caution in the future application of read-across, and may questions the benefit in cases it has been used previously. This relatively data-rich package also exposed where similarity could be drawn if additional endpoints or information were not available, thus creating artificial support for certain endpoints/chemicals being good candidates for read-across. Though this workflow presents transparent biological, chemical, and combined biological/chemical evaluation, moreover it has shown that relatively data-rich, structurally-similar compounds may not lend themselves to strictly biological or chemical grouping. Further showing that even an archetypal dataset fitting read-across criteria requires complex consideration and may not produce a touchstone for utility. These tools allow the integration and visualization of fundamental properties, historical *in vivo* endpoints, as well as novel data types to build an informative, comprehensive approach to integrated chemical classification and read-across categorization.

REFERENCES

Ball, N., M. Bartels, R. Budinsky, J. Klapacz, S. Hays, et al. (2014). "The challenge of using read-across within the EU REACH regulatory framework; how much uncertainty is too much? Dipropylene glycol methyl ether acetate, an exemplary case study." Regul Toxicol Pharmacol 68(2): 212-221.

Berggren, E., P. Amcoff, R. Benigni, K. Blackburn, E. Carney, et al. (2015). "Chemical Safety Assessment Using Read-Across: Assessing the Use of Novel Testing Methods to Strengthen the Evidence Base for Decision Making." Environ Health Perspect 123(12): 1232-1240.

The Dow Chemical Company (2013). Glycol ethers P-series category. Unpublished Internal Document. The Dow Chemical Company, Midland, MI.

European Centre for Ecotoxicology and Toxicology of Chemicals (2005). The Toxicology of Glycol Ethers and its Relevance to Man. http://www.ecetoc.org/publication/tr-095-vol-i-the-toxicology-of-glycol-ethers-and-its-relevance-to-man-fourth-edition-volume-i/.

European Chemicals Agency (2011). "The Use of Alternatives to Testing on Animals for the REACH Regulation." http://echa.europa.eu/documents/10162/13639/alternatives_test_animals_2011_en.pdf.

European Chemicals Agency (2013). Grouping of substances and read-across approach – an illustrative example. Helsinki, Finland, European Chemicals Agency. http://echa.europa.eu/documents/10162/13628/read_across_introductory_note_en.pdf

European Chemicals Agency (2015). Read-Across Assessment Framework (RAAF). Helsinki, Finland, European Chemicals Agency. http://echa.europa.eu/documents/10162/13628/raaf_en.pdf

H.R.2576, 2016. Frank R. Lautenberg Chemical Safety for the 21st Century Act. 114th Congress. https://www.congress.gov/bill/114th-congress/house-bill/2576

Hewitt, M., C.M. Ellison, S.J. Enoch, J.C. Madden and M.T. Cronin (2010). "Integrating (Q)SAR models, expert systems and read-across approaches for the prediction of developmental toxicity." Reprod Toxicol 30(1): 147-160.

Linkov, I., O. Massey, J. Keisler, I. Rusyn and T. Hartung (2015). "From "weight of evidence" to quantitative data integration using multicriteria decision analysis and Bayesian methods." ALTEX 32(1): 3-8.

Low, Y., A. Sedykh, D. Fourches, A. Golbraikh, M. Whelan, et al. (2013). "Integrative chemical-biological read-across approach for chemical hazard classification." Chem Res Toxicol 26(8): 1199-1208.

Low, Y.S., A.Y. Sedykh, I. Rusyn and A. Tropsha (2014). "Integrative approaches for predicting in vivo effects of chemicals from their structural descriptors and the results of short-term biological assays." Curr Top Med Chem 14(11): 1356-1364.

National Research Council (2014). A framework to guide selection of chemical alternatives. Washington, DC, The National Academies Press.

National Research Council (2015). Application of modern toxicology approaches for predicting acute toxicity for chemical defense. Washington, DC, The National Academies Press.

Niska, H., K. Tuppurainen, J.P. Skon, A.K. Mallett and M. Kolehmainen (2008). "Characterisation of the chemical and biological properties of molecules with QSAR/QSPR and chemical grouping, and its application to a group of alkyl ethers." SAR QSAR Environ Res 19(3-4): 263-284.

Oxygenated Solvent Producers Association (OSPA). (2012). "Glycol Ethers Facts and Figures." from http://www.glycol-ethers.eu/what-are-glycol-ethers/facts-and-figures.

Patlewicz, G., et al., 2013. Workshop: use of "read-across" for chemical safety assessment under REACH. Regul Toxicol Pharmacol. 65, 226-228.

Patlewicz, G., N. Ball, R.A. Becker, E.D. Booth, M.T. Cronin, et al. (2014). "Read-across approaches--misconceptions, promises and challenges ahead." ALTEX 31(4): 387-396.

Patlewicz, G., et al., 2015. Building scientific confidence in the development and evaluation of read-across. Regul Toxicol Pharmacol. 72, 117-133.

R Core Team (2015). R: A language and environment for statistical computing. Vienna, Austria, R Foundation for Statistical Computing.

Reif, D.M., M.T. Martin, S.W. Tan, K.A. Houck, R.S. Judson, et al. (2010). "Endocrine profiling and prioritization of environmental chemicals using ToxCast data." Environ Health Perspect 118(12): 1714-1720.

Reif, D.M., M. Sypa, E.F. Lock, F.A. Wright, A. Wilson, et al. (2013). "ToxPi GUI: an interactive visualization tool for transparent integration of data from diverse sources of evidence." Bioinformatics 29(3): 402-403.

Steinbeck, C., et al., 2003. The Chemistry Development Kit (CDK): an open-source Java library for Chemo- and Bioinformatics. J Chem Inf Comput Sci. 43, 493-500.

Vink, S.R., J. Mikkers, T. Bouwman, H. Marquart and E.D. Kroese (2010). "Use of read-across and tiered exposure assessment in risk assessment under REACH--a case study on a phase-in substance." Regul Toxicol Pharmacol 58(1): 64-71.

Wu, S., et al., 2010. A framework for using structural, reactivity, metabolic and physicochemical similarity to evaluate the suitability of analogs for SAR-based toxicological assessments. Regul.Toxicol Pharmacol. 56**,** 67-81.

APPENDIX

Table A-1. GHS-like transformation criteria. Rescaling criteria applied to the Mammalian Toxicity data. Each category had distinct specified endpoints, as seen in Table 2, and for those endpoint values, the chart below was used to rescale the data into a [0, 1] ToxPi format. (*) Denotes endpoints where the data were available in a Boolean Yes/No format. Additional, non-Boolean, Reproductive endpoints follow the Developmental Toxicity scale.

Toxicity Category	***Transformation Criteria***
Category 1= 1; Category 2 = .75; Category 3 = 0.5; Category 4= 0.25; Category 5=0	
Acute Toxicity	Oral: ≤5 = 1; ≤50 = 0.75; ≤300 = 0.5; ≤2000 = 0.25; >2000 = 0 Inhalation: ≤100 = 1; ≤500 = 0.75; ≤2500 = 0.5; ≤5000 = 0.25; >5000 = 0 Dermal: ≤50 = 1; ≤200 = 0.75; ≤1000 = 0.5; ≤2000 = 0.25; >2000 = 0
Repeat Dose Toxicity	NOEL <30 = 1; <100 = 0.75; <300 = 0.5; <1000 = 0.25; >1000 = 0
Developmental Toxicity	NOEL <30 = 1; <100 = 0.75; <300 = 0.5; <1000 = 0.25; >1000 = 0
Reproductive Toxicity*	Yes=1; No=0
Skin and Eye Irritation:	None=0; Slight=.25; Slight to Moderate=.5; Moderate=.75; Severe=1
Skin Sensitization*	Yes = 1, No = 0