Comparison of Modeling Approaches to Prioritize Chemicals Based on Estimates of Exposure and Exposure Potential

Jade Mitchell[a,1], Jon A. Arnot[b], Olivier Jolliet[c], Panos Georgopolous[d], Sastry Isukapalli[e], Surajit Dasgupta[f], Muhilan Pandian[g], John Wambaugh[h], Peter Egeghy[i], Elaine A. Cohen Hubal[j], and Daniel A. Vallero[k,2]

[a] U.S. Environmental Protection Agency, National Exposure Research Laboratory, Research Triangle Park, NC, 2771; USA
[1] *Present address*: Michigan State University, Department of Biosystems and Agricultural Engineering, 206 Farall Hall, East Lansing, MI 48823; jade@msu.edu
[b] University of Toronto Scarborough, Department of Physical and Environmental Sciences, Toronto, ON, Canada; Jon.Arnot@utoronto.ca
[c] University of Michigan, School of Public Health, Department of Environmental Health Sciences, Ann Arbor, MI, USA; ojolliet@umich.edu
[d] Environmental and Occupational Health Sciences Institute, Piscataway, NJ, USA; panosg@fidelio.rutgers.edu
[e] Environmental and Occupational Health Sciences Institute, Piscataway, NJ, USA; sisukapalli@gmail.com
[f] Versar, Inc., Exposure and Risk Assessment Division, Springfield, VA, USA; SDasgupta@versar.com
[g] Infoscientific, Henderson, NV, USA; muhilan@infoscientific.com
[h] U.S. Environmental Protection Agency, National Center for Computational Toxicology, Research Triangle Park, NC, 2771; USA; wambaugh.john@epa.gov
[I] U.S. Environmental Protection Agency, National Exposure Research Laboratory, Research Triangle Park, NC, 2771; USA; egeghy.peter@epa.gov
[j] U.S. Environmental Protection Agency, National Center for Computational Toxicology, Research Triangle Park, NC, 2771; USA; hubal.elaine@epa.gov
[k, 2] Corresponding author: U.S. Environmental Protection Agency, National Exposure Research Laboratory, Research Triangle Park, NC, 2771; USA email: vallero.daniel@epa.gov; 919-541-3306

ABSTRACT

While only limited data are available to characterize the potential toxicity of over 8 million commercially available chemical substances, there is even less information available on the exposure and use-scenarios that are required to link potential toxicity to human and ecological health outcomes. Recent improvements and advances such as high throughput data gathering, high performance computational capabilities, and predictive chemical inherency methodology make this an opportune time to develop an exposure-based prioritization approach that can systematically utilize and link the asymmetrical bodies of knowledge for hazard and exposure. In response to the US EPA's need to develop novel approaches and tools for rapidly prioritizing chemicals, a "Challenge" was issued to several exposure model developers to aid understanding of current systems in a broader sense and to assist the US EPA's effort to develop an approach comparable to other international efforts. A common set of chemicals were prioritized under each current approach. The results are presented herein along with a comparative analysis of the rankings of the chemicals based on metrics of exposure potential or actual exposure estimates. The analysis illustrates the similarities and differences across the domains of information incorporated in each modeling approach. The overall findings indicate a need to reconcile exposures from diffuse, indirect sources (far-field) with exposures from directly, applied chemicals in consumer products or resulting from the presence of a chemical in a microenvironment like a home or vehicle. Additionally, the exposure scenario, including the mode of entry into the environment (i.e. through air, water or sediment) appears to be an important determinant of the level of agreement between modeling approaches.

1. INTRODUCTION

The U.S. Environmental Protection Agency (EPA) regulates chemical substances based on their potential to cause human health and ecological risk. EPA's risk-assessment practices, which provide the scientific foundation for regulatory decisions, continue to follow the recommendations by the National Research Council (NRC (National Research Council) 1983), calling for scientific rigor in characterizing an agent's hazard, dose-response, exposure and effects.. The paucity of sufficient information to evaluate chemical exposure and effects led to the NRC's evaluation of this process and a call for a more integrated assessment of exposure and toxicity (NRC 2009). The NRC's recommendations directly relates to risks presented from chemical substances found in products and materials used by society.

The primary purpose of the Toxic Substances Control Act (TSCA) is "to assure innovation and commerce in such chemical substances and mixtures do not present an unreasonable risk of injury to health or the environment" (TSCA § 2 (b)(3)). At present, an unprecedented and increasing number of chemicals for which risks must be assessed, at least at some screening level, continue to be added to the TSCA inventory (over 80,000). Approximately 100,000 chemical substances currently exist in commerce worldwide (Muir and Howard 2006). Hence, the urgent need for new, broadly applicable tools to facilitate rapid risk characterization (Dix, Houck et al. 2007) is widely recognized and has been well established in the literature. The U.S. is certainly not alone in this effort, as legislated mandates for efficient risk based screening, categorization, classification and prioritization of chemicals exist in the European Union and Canada as well (Egeghy, Vallero et al. 2011).

High throughput screening models are needed for both hazard (toxicity) and exposure, since risk is a function of both. Current advances in hazard utilize computational chemistry and *in silico*

methods for high-throughput screening (HTS) and various toxicogenomic methodologies. These have accelerated the prediction of potential toxicity for prioritization by providing a greater quantity and diversity of data. Notably, EPA's ToxCast™ program applies a battery of rapid *in vitro* assays to predict toxicity. To complete the high throughput risk screening, however, exposure information must also be acquired in a similarly rapid manner, and compared with results from ToxCast™. Thus, there is a need to rapidly assess chemicals on the basis of 'biologically-relevance' human exposures to target research and improve risk assessments (Cohen Hubal, Richard et al. 2010). In 2009, EPA launched its ExpoCast™ initiative to address this need for exposure data, novel modeling approaches and discriminating metrics to screen and evaluate chemicals based on potential for human exposures [Cohen Hubal, Richard et al. 2010)].

The use and development of exposure models as well as expert judgment have aided in the endeavor to screen and prioritize chemicals when exposure measurements are limited or unavailable (Schinkel et al 2011, Jayjock et al 2009). In its 1995 review of the state-of-the-science, the Society of Environmental Toxicology and Chemistry found that the chemical ranking and scoring systems developed over the previous 20 years were widely diverse in the factors used for screening potential exposure, including: (1) chemical marketing data; (2) emission data; (3) physical-chemical properties; (3) persistence and transformation processes; (4) monitoring data or other measured concentrations; (5) modeled or estimated concentrations; (6) receptor characteristics and exposure setting – including consumer, worker exposure related to frequency, duration, and intensity; and (7) exposure expressed as intake (Swanson and Socha 1997). More recently, Egeghy and coauthors review modeling tools and approaches available for prioritizing manufactured chemicals. These approaches tend to fall into two categories: the first focuses on characterizing the fate and transport of a chemical following release into the environment, while the second focuses on understanding exposures resulting from use and interaction with consumer

products (Egeghy, Vallero et al. 2011). Egeghy and coauthors identified the need to better evaluate both categories of models and approaches to assess strengths and limitations for rapidly evaluating and ranking chemicals based on potential for exposure.

In the context of this manuscript, the term far-field refers to indirect chemical exposures from environmental sources, as in water, air, soil, and food stuffs; and the term near-field refers to exposures in microenvironments such as buildings and cars. The near-field includes direct exposures, e.g., as application and use of consumer products and Personal Care Products (PCPs) and indirect exposures to ambient sources e.g., off-gassing of building materials, consumer products, dust ingestion (Jayjock, Chaisson et al. 2009).

An initial study to leverage existing tools is presented herein. We compare and evaluate the capabilities of existing tools and identify the gaps which must be addressed to develop an exposure-based prioritization approach that can be applied to rapidly and efficiently evaluate broad classes of chemicals. The purpose of this paper is to present a comparative analysis of exposure-based prioritization results for a common set of chemicals using several different modeling approaches and exposure metrics. This analysis is designed to address how consistent the models' rankings are with each other and if the models consistently rank the same chemicals higher or lower than others.

2. METHODS

An exposure based prioritization model challenge was issued publicly in 2010 (http://www.epa.gov/ncct/expocast/exposure_based_challenge.html) to elicit the experience of developers of existing prioritization schemes or of models for rapid estimation of exposure potential that can be used to inform exposure-based prioritization of chemicals. Challenge participants were charged with using existing approaches to prioritize a sets of compounds based

on potential for exposure. The objective of the challenge was to gain a better understanding of the process by which existing approaches evaluate potential exposures. As such, the main interest was in an explicit and transparent documentation of each approach. Employing existing tools further facilitates an analysis identifying the type, quantity and quality of information needed for a more comprehensive approach toward exposure-based prioritization.

2.1 QUANTITATIVE EXPOSURE METRICS

Metrics for assessing chemical exposure and exposure potential referred to in the present study are the intake rate (*iR*; *e.g.,* μg/d or kg/h), the intake fraction (*iF*; *e.g.,* kg-intake/kg-emission), and the concentration in an organism such as a human (*C*; *e.g.,* μg/kg). If desired, the intake rate can also be calculated on a body weight basis (iR_{BW}; *e.g.,* μg kg-bw^{-1}day^{-1}). These metrics can be calculated using a consistent, arbitrary "unit emission" rate (E_U; *e.g.,* kg/h) for all chemicals to screen, compare and prioritize chemicals based on relative exposure potential, *i.e.,* using *iF* or C_U. The steady state intake rate based on a unit emission rate (iR_U; kg/h) provides similar screening and ranking information as the steady state intake fraction *iF* since the intake fraction is the chemical intake rate normalized to the emission rate, *i.e.,* $iF = iR/E$. The *iF* can be calculated on an individual basis, an age-class specific basis (iF_{AC}, e.g. "toddlers" vs. "adults"), or for a human population in a defined spatial region (iF_{POP}). The latter two *iF* endpoints are thus dependent on population and demographic information. Intake rates and intake fractions can include the sum of all exposure routes (*i.e.,* aggregate exposures) or they can be calculated for specific sources (*i.e.,* for air, water, and food stuffs individually) to compare the relative importance of different exposure routes and to identify those routes of exposure that are expected to be highest for a particular chemical. Relative exposure potential metrics are thus independent of an actual emission rate and are useful for chemical screening evaluations due to the substantial uncertainty in actual emission rate estimates. Relative exposure potential comparisons are a function of the chemical and

environmental properties (persistence, bioaccumulation) and the underlying assumptions and parameters used to characterize the environmental and human conditions in the models.

Alternatively, exposure metrics can be calculated using estimates of "actual emission" rates (E_A; *e.g.,* kg/h), thus providing indices for screening and priority setting based on actual exposure estimates, i.e., C_A or iR_A. Estimates of actual chemical exposure using concentrations or intake rates are directly applicable in risk based chemical assessments by comparing body/tissue concentrations or intake rates of exposures with those concentrations or rates of intake associated with effect or no effect levels. Unlike iR_A and iF, C is an internal dose metric and therefore it depends on absorption (*i.e.* including gastrointestinal biomagnification from dietary exposures) into the body and elimination processes such as fecal egestion, urinary excretion, respiratory exhalation, and importantly for most chemicals, metabolic biotransformation.

2.2 SEMI-QUANTITATIVE METRICS

A set of semi-quantitative measures of potential exposures to the chemical of concern are also presented as a part of a tiered screening approach. These metrics are based on a combination of available quantitative information on releases and concentrations, qualitative information on types and degree of exposures reported in the literature, and expert judgment on various facets of the exposures. The four population-based metrics considered for exposure based ranking are: pervasiveness, persistence, severity, and efficacy. The semi-quantitative metrics reflect: (i) how widespread the exposures could be within the general US population (pervasiveness); (ii) the temporal frequency and/or duration of such exposures (persistence); (iii) the potential for high levels of such exposures (severity); and (iv) the potential of the contact with the chemical to result in intake/uptake (efficacy).

2.3 CHEMICALS

A set of 52 chemicals (or defined mixtures) was provided (Table 1). These chemicals are representative of several broad categories in terms of physical-chemical properties and typical

intended use. Thus, exposure to these substances can be assessed across multiple routes and pathways depending on the prioritization model or approach. The majority of the set is comprised of chemicals known to have a relative abundance of publically available exposure-related data accessible for modeling; conducting this pilot from a 'data rich' perspective was intended to limit the burden of data collection. Several chemicals with meager available exposure data, however, were also included to test the capabilities of the models for addressing the vast majority of chemicals in wide commercial use with little or no existing exposure-related information (Egeghy et al., 2011b).

2.4 MODELS USED

Four model developers responded to the exposure based prioritization challenge allowing for the investigation of five models:

1. Far-field Human Exposure (FHX) (Arnot, Mackay et al. 2010)
2. Risk Assessment IDentification And Ranking(RAIDAR) (Arnot, Mackay et al. 2006)
3. UNEP-SETAC Toxicity Model (USEtox™™) (Rosenbaum, Bachmann et al. 2008)
4. GExFRAME/Scibin (Kephalopoulos, Arvanitis et al. 2008)
5. Prioritization/Ranking of Toxic Exposures with GIS Extension (PRoTEGE) – derived from Modeling ENvironment for Total Risk studies (MENTOR) (Georgopoulos and Lioy 2006)

Additionally, EPA included two of its own currently used systems: (1) the Exposure and Fate Assessment Screening Tool (EFASTv2) (http://www.epa.gov/opptintr/exposure/pubs/efast.htm) and (2) the Stochastic Human Exposure and Dose Simulation (SHEDS) model (http://www.epa.gov/heasd/products/sheds_multimedia/sheds_mm.html), Some distinguishing features of the models were recently described by Egeghy et al. (2011).

2.5 REQUESTED OUTPUTS

In addition to providing ranking results for the chemicals, the modelers were also asked to transparently describe the following characteristics of the modeling approaches:

- Protocol for applying prioritization tool/approach
- Model structure, algorithms, and assumptions
- Target sources (near and far field as defined above)
- Target population of interest
- Emission characteristics
- Exposure pathways (media, routes)
- Use of professional/expert judgment
- Defaults (exposure factors, etc.)
- Consumer product use categories and selection of sentinel products
- Selection of exposure scenarios (inputs to scenarios)
- Basis for prioritization score
- Units of prioritization metric and other model outputs

2.6 COMPARISON OF PRIORITIZATION RESULTS

For the models which produced quantitative metrics (RAIDAR, FHX, USEtox, Protégé, and EFAST), a quantitative comparison was conducted based on each chemical's relative ranking. This initial comparison also provides a basis for further investigation where significant differences between modeling results would warrant focusing on the differences in model algorithms, defaults, data input sources, etc.

Each model developer identified at least one exposure metric for use as the basis for prioritization; however, the metrics from the different approaches were not always directly comparable. Some of the endpoints used in the comparisons below are "exposure potential" i.e., intake fractions and some are "exposure" endpoints, i.e., internal concentrations or body burdens. (For more information on exposure metric development the reader is directed to Rosenbaum et al 2008, Bennett et al 2002, and Arnot et al 2010.) The modeling approaches also used different model input parameters for chemical properties, use and release rates. Based on the exposure metrics, the chemicals were ranked according to magnitude thereby creating ordinal values from the measured or modeled values with monotonic differences between chemical exposure ranks. While this is a convenient way to view relative rankings between chemicals we note that with this approach some information is lost because of the transformation. For example, multiple chemicals may have fractional differences in the value of the exposure metric, but once transformed these differences

become exaggerated by the ordinal scale. For the purpose of prioritization, the utility of the lost information may be considered negligible as it is not the intent of these applications to produce precise measures of exposure. The actual values of the exposure metrics were however used to assess the degree of differentiation between chemicals in results from each approach. This is of interest especially when uncertainty around the estimates is considered. For example, if many chemicals have the same exposure estimates, which is a possible artifact of using default values for missing data, the ability of the approach to separate high, low and moderate exposure potential will be impaired. This topic is included in the qualitative discussion of the results. Table 2 summarizes the comparisons conducted in this analysis based on the compatibility of the exposure metrics provided.

The quantitative analysis consisted of four methods:

(i) Correlation between the ranks of each chemical produced by the models was assessed using two non-parametric measures of rank correlation, Spearman rho and Kendall tau. The Spearman rank correlation indicates the direction of association between two variables that can be related by any monotonic function. The Kendall tau correlation depends on the ratio of concordant pairs to discordant pairs or the number of inversions needed to transform one rank order into the other. The Kendall tau coefficient can be interpreted probabilistically as the difference between the probability of the set of ranked objects being in the same order and the probability of the ranked objects being in a different order. (Abdi 2007) While in most cases, the Spearman rho produces a higher correlation coefficient than the Kendall tau, the two measures can yield meaningfully different results. Spearman's rho is more sensitive to a few large deviations than Kendall's tau. This information can be useful if one wants to consider using several modeling approaches in combination to form a consensus ranking for each chemical in the future; therefore both correlation measures were evaluated. Chemicals

not ranked by a particular model were eliminated from the correlation analyses and the remaining chemicals were re-ranked relatively.

(ii) The chemicals were sorted into 5 equally distributed bins based on their ranks. Correlation among the bin designations was assessed along with the expert judgment driven semi-quantitative metrics of exposures from PRoTEGE.

(iii) To evaluate consistency among the modeling approaches without considering the source of variability contributed by the chemical being modeled, a randomized block design was applied to the data blocking on the specific chemical. The Friedman test, a non-parametric test similar to the parametric repeated measures ANOVA, was then conducted to detect differences across each modeling approach. The Friedman test is the measure of the aggregate degree to which each modeling approach differs and is used to compare three or more paired groupings rather than two pairwise groups described by the correlation coefficients. The hypotheses for the comparison across the repeated measures are: The null hypothesis (H0) - The distributions are the same across repeated measures; and the alternative hypothesis (H1) -The distributions across repeated measures are different.

The test statistic for the Friedman's test is a Chi-square with [(number of repeated measures)-1] degrees of freedom.

Tables 3a and 3b describe the data selected for comparison.

3. RESULTS

3.1 COMPARISON OF EVALUATION SCHEMES/MODELS

We briefly summarize a qualitative comparison of the aforementioned model characteristics to provide context for the prioritization results, though this study focuses on the metrics utilized for prioritization (exposure estimates and rankings) from each approach without regard to the

underlying purpose, algorithm and assumptions involved in each. RAIDAR, FHX and USEtox are far-field models representing exposure from diffuse sources to the general human population of different regional spatial scales. These three models use mechanistic mass balance approaches to simulate fate and transport in different environmental compartments. Models for bioaccumulation into food sources destined for human consumption are also included in the exposure pathways. The intake fraction or mass of substance available for contact with an organism per mass emitted to the environment was identified as the primary metric for exposure based prioritization from these models. Additionally, intake rate from the RAIDAR and FHX models can be used. The FHX model provides age class specific rankings for chemicals, but these were not significantly different than the rankings for the adult age class so they were not included in our comparative analysis. Additionally only the RAIDAR model includes absorption, distribution, metabolism and excretion (ADME) processes and an internal concentration in humans as a metric for prioritization. While this metric may be considered most 'biologically relevant' and compatible with toxicity data some technical issues have been noted elsewhere in regards to corroborating this metric against measured biomarkers (Georgopoulos, et al. 2009). For simplicity, each of the far field models can be run based on an equivalent unit emission and equilibrium assumption for each environmental compartment to provide a relative estimate for each chemical or with more advanced fate assumptions. This approach may be useful for making comparisons among broad ranges of chemicals for which actual emission data is unavailable though this contributes more uncertainty in the chemical screening. Under the unit emission assumption, RAIDAR and FHX provided rankings for 45 chemicals and USEtox provided rankings for 42 chemicals. When actual US emission estimates were possible, USEtox estimated exposures for 29 chemicals and RAIDAR estimated exposures for 10 chemicals (though actual EU emission rates were available for the remainder of the chemicals assessed by RAIDAR).

Intake is an intermediary variable in USEtox to be coupled with toxicity to form a combined exposure/toxicity "damage" endpoint. The intake value from USEtox alone is not considered a suitable metric for ranking but provides a means of comparison on the basis of exposure. The three far-field models can be parameterized with chemical specific information from measurement databases or estimated from readily available Quantitative Structural Activity Relationship (QSAR) models.

The EFAST model provides similar far-field modeling estimate, but also considered risk management interventions to reduce exposure. It also includes a consumer product module. EFAST2 is a screening-level computer tool that allows users to generate exposure estimates for humans and the environment through various release and exposure scenarios. The chemical-specific input parameters (mostly physicochemical properties) used in the EFAST2 model include: 1) bioconcentration factor (BCF; L/kg), 2) concentration of concern (CoC; ppb), 3) wastewater treatment removal rate (WWT; %), 4) incineration removal rate (%), 5) fugitive removal rate (%), 6) ground water migration rate, 7) molecular weight (g/mol), 8) vapor pressure (mm Hg), and 9) weight fraction (%). The modeling endpoints estimated with EFAST2's General Population and Ecological Exposure from Industrial Releases module include the human Acute Dose Rate (ADR in mg/kg-day) and LADD (Lifetime Average Daily Dose (LADD in mg/kg-day) for exposures from drinking water (via surface water releases), groundwater ingestion (via landfill releases), fish ingestion (via surface water releases and subsequent bioconcentration), and inhalation (from stack and fugitive air releases).

The far-field modeling approaches produced the most rapid results of those evaluated in this study, primarily because of the ease of parameterization using available chemical specific properties.

3.1.1 RAIDAR

RAIDAR was used to estimate exposure for X number of challenge chemicals. Figure 1 shows the unit emission based chemical results produced by the RAIDAR model on the basis of individual intake fraction (IF) and internal concentration (Cu) at Tier 1Level II fate model calculations. These assumptions do not require a model of entry into the environment as they assume instantaneous equilibrium in all physical compartments. The results in Figure 1 highlight the increase in differentiation among chemicals when ADME processes are included (~12 orders of magnitude when included vs. ~6 orders when not included). ADME processes made a significant difference in relative rankings for those chemicals predicted to have moderate exposures levels. For those predicted to have either very low or very high exposure levels, results from both internal and external exposure metrics were highly correlated. In RAIDAR this difference was a more significant factor in differentiating potential exposures than the assumed emission release compartment (i.e. air, water or soil) which is unknown for many chemicals recommendations comprehensive

3.1.2 USEtox

USEtox was used to provide exposure predictions for 45 number of challenge chemicals. An example of the results of predicted population intake fraction from the USEtox model is provided in Figure 2. This approach produced differentiation among chemicals in the same order of magnitude as RAIDAR, (i.e. the range covered over seven orders of magnitude.) The USEtox model exposure framework was recently expanded to include indoor air as an additional emission compartment (Hellweg et al., 2009) and further developed to account for sorption to indoor walls as a removal pathway (Wenger et al., 2012). Because human exposure to manufactured chemicals in consumer products is of concern, an indoor compartment will improve relevance of screening-level model predictions.

3.1.3 EFAST

The EFAST2 model produced results for 48 chemicals for water releases only. Each exposure scenario (drinking water, fish ingestion, and groundwater/landfill routes) was ranked separately and the chemicals were assigned a relative score (1-4) based on the distribution of their dose estimates For comparison with USEtox and RAIDAR the dose estimates for emission to water were combined and each chemical was ordinally ranked. Since the model was developed for screening level estimates, the results are highly conservative and rely on many default values. In comparison with USEtox and RAIDAR the temporal scale of the release conditions also differs in EFAST2. Where USEtox and RAIDAR assume steady state releases, the EFAST2 model can evaluate exposures from several different temporal release patterns. For this evaluation a unit emission was considered for a twenty day period to produce exposures measured in ADR and LADD as described above. Because of the assumptions used in this relative analysis of chemicals, both the ADR and LADD estimates were perfectly correlated, so only the LADD values are used in the following comparative analysis. An example of the EFAST2 results for the total exposure from water releases is illustrated in Figure 3. Differentiation of estimates spanned 4 orders of magnitude. Additional results from EFAST included 9 chemicals ranked using the consumer module. These estimates are calculated solely from inhalation (no other route) using given chemical-specific inputs of molecular weight and vapor pressure. Only 11 chemicals were identified to have potential household consumer applications (that fit within the EFAST model), two chemicals were excluded because they did not include an inhalation assessment (this is the case when only the bar soap scenario applied). The consumer use scenarios applied to the chemicals assessed include: (1) general purpose cleaner; (2) bar soap; (3) laundry detergent; and (4) latex paint.

3.1.4 GExFRAME

GExFRAME and PRoTEGE provided rankings for microenvironmental exposures, though both used different approaches. PRoTEGE is based on a more sophisticated predecessor, MENTOR (Modeling ENvironment for TOtal Risk) and estimates both near-field and far-field exposures. MENTOR

supports detailed person-oriented source-to-dose exposure modeling for mixtures of multimedia contaminants, allowing a focus on specific locations and subpopulations. Neither MENTOR nor GExFRAME are considered single models. Both are modeling systems or modeling environments capable of accommodating various algorithms to integrate scientific data and models for estimating exposures. For this analysis, the CARES (Cumulative and Aggregate Risk Evaluation System) program was integrated with GExFRAME to assess dietary and non-dietary residential exposures. The results from the use of GExFRAME in this exercise were purely qualitative and could not be included in the quantitative analysis. Chemicals used in consumer products were grouped into scenarios that defined similar source characteristics, media dispersion characteristics and exposure pathways. Based on these categories a scenario specific set of exposure algorithms were assigned. A default set of inputs are available for conducting exposure assessments within each set of algorithms. The 52 challenge chemicals were classified as described in Table 4.

The GExFrame analysis produced only categorical values because GExFrame requires measured or monitored data in near field exposure media, e.g. breathable air, contact surfaces, ingestible soil, dust in air and surfaces, food and drinking water and other near field information. That is, GExFrame is designed to provide a high tier, low throughput chemical characterization. Chemicals are characterized, but not specifically prioritized. Chemicals that fall within Categories 1 and 2, however, are likely to exhibit relatively low exposure potential. By extension, chemicals in Categories 3, 4, and 5 may well exhibit increased exposure potential and, consequently, such chemicals pose important risk. Finally, chemicals in Category 6 are regulated under the Federal Insecticide, Fungicide, and Rodenticide Act and have undergone higher tier exposure assessment. Exposure potential characterization is also use for prioritizing chemicals in terms of need for further testing. Thus, chemicals in Category 6, notwithstanding their high exposure, may be

considered lower priority because much is known about the chemical class in general and higher tiered models are available to produce more accurate results.

3.1.4 SHEDS

The US EPA SHEDS model was determined to be useful for screening chemicals when specific use categories and scenarios could be established. Parameterization of the model could be conducted using default values, but this was not the intended purpose of such high tier models so it presented some challenges. Future work to modify the SHEDS model for lower tiered assessments is required. Five different use categories were gleaned from the list of chemicals in this analysis: industrial/occupational additives & by products, plastics, commercial additives, pesticides/herbicides and natural risks (i.e. arsenic, lead, manganese, cadmium). Various combinations of SHEDS-Dietary, SHEDS-Soil/Dust, SHEDS-Residential, SHEDS-CCA (chromate copper arsenate, formerly SHEDS Wood) , APEX (Air Pollutants Exposure Model) and SHEDS-Air Toxics were determined to be useful for all of these categories except for the plastics if appropriately modified with default scenarios.

3.1.5 PRoTEGE

The PRoTEGE system facilitates screening level exposure calculations at multiple tiers, utilizing available data on

- Chemical production and usage,
- Environmental releases,
- Environmental concentrations in multiple media and microenvironments, and
- Age- and gender-specific population distributions of major physiological and behavioral attributes.

PRoTEGE can estimate exposures during chemical manufacturing, chemical transportation, product manufacturing, product use, and product disposal or as a result of intentional or unintentional

release of the chemical in the environment and its subsequent contact with the individual through one or more exposure routes (inhalation, ingestion, dermal contact). A summary of these results are illustrated in Figure 4.

"Tier 2" exposure metrics of PRoTEGE provide quantitative measures of potential exposures to the chemicals of concern using probability distributions of multimedia contaminant concentrations, combined with distributions of physiological and behavioral factors. These metrics are primarily based on available nationwide data and are summarized in Figure 5. The exposure estimates spanned 6 to 10 orders of magnitude differentiating among the chemicals evaluated before they were transformed into categories. Time to obtain these chemical specific data points may be excessive when applied to larger sets of chemicals for screening and prioritization purposes.

Results from PRoTEGE were provided for 55 chemicals in Tier 1 – those listed in Table 1 and additional perchlorate salts - sodium perchlorate, potassium perchlorate, and magnesium perchlorate. In Tier 2, estimates were obtained for 47 chemicals – both median values and 95th percentile estimates were provided.

3.2 COMPARISON OF PRIORITIZATION RESULTS

3.2.1. Actual Emissions

Comparison between the prioritization results were only made where they were deemed appropriate as described in Tables 2, 3a and 3b – with similar emission assumptions, modes of entry, etc . While RAIDAR and USEtox mainly provide estimates of exposures due to environmental releases versus exposure due to indoor sources or consumer product use exposure estimates, PRoTEGE estimates are equivalent to intakes for the general population and therefore provide a means of comparison with RAIDAR and USEtox. In other words, although the model structures are different, the exposure metric is similar and therefore these three models can be compared. In

Figure 6, the exposure metrics across all pathways based on actual emission estimates were used to rank each chemical for each modeling approach. This figure shows a side by side comparison of the magnitude of the ranking results based on color. Red denotes the highest potential exposure and green denotes the lowest potential exposure. Numerically, a value of 1 corresponds to the highest exposure and 55 correspond to the lowest exposure. White spaces indicate no value as there were chemicals for which an approach failed to produce a result. As expected, similarities between the median and 95th percentile estimates from RAIDAR were observed for the majority of chemicals. PRoTEGE showed more deviations between median estimates and upper bound estimates than RAIDAR. The PRoTEGE distributions include both variability and uncertainty in concentrations as well as demographic variability among physiological and behavioral factors. There are several cases where USEtox generated comparable rankings with PRoTEGE (e.g. 2,4-D, ethane, 1,1,2,2-tetrachloro- , methoxychlor and pentachlorophenol) and several cases where RAIDAR and PRoTEGE produce results at opposite extremes of exposure likelihood e.g. Aroclor_1260, Aroclor_1260, aldicarb, hexachlorobenzene, octaBDE pentachlorophenol, etc.). It should be noted again that the RAIDAR results use EU actual emission estimates for most chemicals. There were only four chemicals across these 3 models where RAIDAR also used US data. These comparisons are shown in Table 5. When ranked relatively, USEtox and PRoTEGE produce the same results for median estimates. Two of the four chemicals are ranked the same using RAIDAR and PRoTEGE when the 95th percentile values are compared.

A total of 18 chemicals could be compared across USEtox and PRoTEGE with actual emission estimates. The Spearman rho correlation coefficient is 0.81 and the Kendall tau is 0.63. When USEtox is compared with the 95th percentile estimates from PRoTEGE similar results were obtained (0.82 and 0.65, respectively). These values represent very high agreement between the two approaches.

484

485 A second analysis to evaluate the modeling approaches using actual emission rates was conducted
486 by comparing the ability of each approach to bin chemicals. There were 16 chemicals included in
487 this analysis. Figure 7 summarizes the Spearman rho correlation coefficients obtained when the
488 chemicals were sorted into 5 bins. The Kendall tau values were consistently lower for all values
489 reported so they are not included. Qualitative metrics from the PRoTEGE model were included in
490 this analysis. As seen by the darker green shading, the RAIDAR metrics of exposure show no
491 positive association with the exposure estimates from PRoTEGE or USEtox. Interestingly, the
492 RAIDAR results are positively correlated with the inhalation, dermal and aggregate qualitative
493 metrics from Protégé. As expected, the binned quantitative metrics from PRoTEGE and USEtox are
494 inversely proportional with the PRoTEGE qualitative metrics except for ingestion. Within the Tier I
495 qualitative metrics, the ingestion pathway is inversely related to the dermal and inhalation
496 pathways along with the aggregate or dominant pathways which are influenced by these measures.

3.2.2 Unit Emissions

Due to lack of actual emission data for a large number of chemicals, more compounds could be evaluated using a unit emission rate. RAIDAR, USEtox and EFAST produced rankings based on unit emissions to water for 38 chemicals. The Spearman rho correlation is reported in Figure 8. Positive correlation was observed between all measures. RAIDAR and USEtox intake fractions were highly associated with a Spearman rho of 0.7. EFAST2 produced results that were associated with both USEtox and RAIDAR to a lesser extent (Spearman rho = 0.3). The major difference between the later approach and the former approaches is the inclusion of chemical specific removal rates from environmental media. The Kendall tau values were also consistently lower so they are not reported.

A strong association was also revealed between the intake fraction and internal concentration metrics from RAIDAR. This indicates that the inclusion of ADME may be unnecessary at the prioritization stage. However it should be noted that most of the chemicals here are "legacy pollutants" (i.e. persistent). A different pattern may exist for the larger domain of chemicals requiring evaluation.

When rankings based on air emissions were compared (Figure 9), very high association was observed between RAIDAR and USEtox. Comparison of the intake fractions for 38 chemicals resulted in a Spearman rho of 0.9.

To avoid bias from the gross aggregation of source compartments that may affect the rankings, intake fractions for ingestion and inhalation of air and water were disaggregated for comparison. Areas of departure between far-field models (USEtox in contrast with RAIDAR and FHX) are the result of differences in (i) model input parameters (model users did not use consistent input parameters)), (ii) treatment of release (emission) rates to population densities, and model

structure (e.g., differences in fate and food web calculations as well as human contact rates (i.e. diets).Only a comparison between models with consistent model input parameters established for each can shed light to the cause of the observations here , but it useful to see which domains of information and metrics yield similar results.

The Friedman test compares three or more paired groups. The test allowed for comparison between model outputs while separating sources variability. The two major sources of variability are the individual chemical being modeled and the model used for prioritization (e.g. RAIDAR, USEtox, Protégé, EFAST, etc.). Based on the total of 38 chemicals evaluated by each model using the unit emission rate, this Friedman test can identify consistency among approaches aggregately rather than pairwise. This analysis answers the question “Are the models’ rankings consistent with each other? “ or “Are some models ranking chemicals consistently higher or lower than other. In every case the null hypothesis, equal treatment by each model, could not be rejected meaning that there is some consistency between modeling results. The results are summarized in Table 6. While the actual ranking scores of each chemical are dissimilar (as indicated through correlation tests), consistency exists between approaches in a broader sense based in the Friedman test which may indicate that it is appropriate to combine approaches to form a consensus ranking.

4. Discussion

Overall the statistical comparisons reveal that between the far-field models viewed here, RAIDAR, FHX and USEtox, there is close agreement between chemical rankings when the emission compartments are consistent (i.e. water and air). The case study results from these models are dependent on the emissions rates as drivers of the modeling results so expected differences were observed when these inputs were varied from the unit emission assumption. When compared to near field models (PRoTEGE) a significant indication of agreement between rankings exist when the

emission estimates are regionally compatible (i.e. U.S. release inventories). Unfortunately, these actual release quantities are unavailable for a large number of chemicals produced at lower production volumes. Inverse relationships between the qualitative estimates informed in part by expert judgment and the quantitative modeling results were observed to be sensitive to exposure pathway. This is an indication that scenario specific chemical rankings may be important going forward. Another finding to support the need to resolve scenario specific exposure prioritization issues is that within the USEtox modeling results, the rankings of chemicals from the indoor air pathway and outdoor air pathway had a very low inverse Spearman correlation coefficient of -0.18 as well. Time activity data has shown that the population spends most of their time indoors including residential and occupational settings. Therefore the need to include exposure predictions for near-field exposure is significant. Models like GExFRAME and SHEDS are designed to make these scenario dependent predictions, however they are incapable of producing results for a large number of chemicals lacking the sufficient input data for higher tiered assessments. It is clear that characterizing important factors like habits and practices of consumers which drive exposure potential is a critical need. The PRoTEGE model allows for lower tier predictions to capture these factors in a "data poor" environment with semi-qualitative metrics. Where more data are available, median higher tier predictions of exposure potential based on quantitative assessments showed moderate agreement with the RAIDAR model but inverse association with USEtox and Protégé. While these types of metrics offer a promising alternative in the absence of data, it will be important to reconcile the differences between the semi-quantitative and modeled results in terms of specific exposure pathways which appear to be the most significant contributor to the observed differences.

Ultimately the reliability and utility of these approaches is dependent on their ability to rapidly assess thousands of chemicals for which little exposure information is anticipated. Developing a set

of criteria or specific needs from a high throughput exposure based prioritization approach is a necessary precursor. The objectives of such an approach are needed to more clearly define how to balance the tradeoffs between producing rapid results with available information and meeting an acceptable level of confidence in screening. For example, the categorization of chemicals seen here for GExFRAME and SHEDS could be useful in eliminating chemicals from an inventory because of little concern over potential exposure though they do not produce individual exposure estimates. Additionally, these approaches may be parameterized with defaults to shed light on the relevance or potential exposure associated with particular scenarios. The developers of EFAST2 and PRoTEGE, approaches which did produce exposure estimates recommended binning the results into 4 or 5 risk categories based on associated confidence. The term prioritization infers the need to have chemical specific ordinal results but the ability to 'screen for prioritization' should not be necessarily discounted. How comprehensively exposure across the source to receptor continuum is characterized is another issue that needs to be resolved. The far-field models were more inclusive of processes relevant to the entire continuum (though they were more rigorous in processes from source to concentration). The near field models rely on environmental concentrations from scenario specific use of chemicals and focus more on the successive processes from concentration to receptor. Because the entire source to dose continuum is deemed relevant in producing compatible results with high throughput toxicity testing how the results from far-field and near field source models can be combined warrants further exploration, especially because inverse associations were observed in some cases. A further consideration is whether a consensus approach should be developed or an approach that synthesizes the results from models across the source to dose continuum. Equal treatment of chemicals by each model demonstrated by the Friedman test may inform the appropriateness of consensus building when exposure source and pathway considerations are reconciled.

As highlighted by the RAIDAR results, the incorporation of internal exposure in prioritization schemes acknowledging the complexity of exposure is an important topic for exploration. For example, the rankings based on intake fraction vs. internal dose from the challenge results, show that incorporating ADME into prioritization. In other words, in many cases the precise pathway of exposure becomes relatively less important in determining tissue concentrations compared to the efficiency of absorption and the residence time in the body.

The authors recognize some additional factors for consideration in developing exposure based prioritization approaches that fell outside of the scope of modeling results herein. These issues are bulleted below:

- Consideration of chemicals as single substances rather than embedded in products
    a. The issue of confidentiality of individual product usage data is a problem.
    b. There is a separation in the product chain between the companies that make the chemical and the companies making the final products.
    c. Those reporting production volume for a chemical are often unaware of how it may be used downstream.
    d. Decisions on how chemicals may be used are sometimes made by many different smaller companies (formulators) rather than by large manufacturers who are used to working with regulators to provide information for exposure assessment.
- The amount of the chemical released vs. the amount used vs. the amount produced may lead to different prioritization results.
- In terms of modeling, an extension of exposure characterization beyond occurrence, persistence and bioaccumulation is necessary for prioritization of large numbers of chemicals.
- Consideration of the potential lifetime of the person and the products
- Should the function of the chemical or types of products it is used in (toys, paints, etc.) be the basis for exploring exposure scenarios?

5. Conclusions

Currently available approaches show promise for prioritizing chemical ingredients in products according to their potential for exposure, but several gaps in knowledge exist. Chief among these gaps are the paucity of information needed for reliable estimates of exposures for direct and micro-

environmental scenarios and the need for improved understanding of product use and resulting release rates. A next step will be for EPA to simultaneously, evaluate exposure-based prioritization linkages to hazard based chemical prioritization approaches. Since risk is a function of exposure and hazard, such integration is needed for risk-based chemical prioritization. This is consistent with current risk assessment recommendations (notably, by the National Research Council, 2009).

Another need is the ability to extrapolate exposure characterization and ranking from the small subset of data-rich chemicals. This will rely on QSAR and other methods to extrapolate from known chemicals with estimated properties to other chemicals. Some older techniques can be found in EPI Suite, ChemSTEER and similar chemical databases and systems, but updated fate simulation capabilities would have a positive impact on the outcome of future analyses. Uncertainty in using estimated properties for exposure and risk assessment model inputs is expected to be substantial; thus methods to address this uncertainty need to be considered.

Formal techniques in decision making under high uncertainty have greatly improved over the past decade. Although their usefulness was recognized as early as in 1995 in SETAC's state of the science publication (Swanson and Socha 1997), decision analysis techniques had rarely been used in chemical ranking and scoring systems. Utilizing outranking approach in multi-criteria decision analysis has several advantages to exposure based prioritization which are being explored. The model structure allows for formal value of information analysis which addresses the question of what data are most needed to make a prioritization decision allowing EPA to prioritize research needs and request the right information from industry during the pre-manufacture process. Other decision science techniques allow for the modeling of expert judgment when measured or monitored data are absent. Bayesian networks can be part-mechanistic and part-statistical frameworks for incorporating and combining information from data and other sources (i.e. expert

judgment). These approaches are promising for synthesizing important exposure related criteria. Subsequent exercises have been designed to apply current approaches to larger sets of chemicals lacking sufficient data to further facilitate the evaluation of the utility and reliability of these approaches in a heuristic method (Wambaugh et al., submitted).

DISCLAIMER:

The views expressed in this manuscript are those of the authors and do not necessarily reflect the views or policies of the U.S. Environmental Protection Agency.

REFERENCES

Abdi, H. (2007) “The Kendall Rank Correlation Coefficient.” Encyclopedia of Measurement and Statistics. Sage. Thousand Oaks, CA

Arnot, J. A., D. Mackay, et al. (2010). "Estimating farfield organic chemical exposures, intake rates and intake fractions to human age classes." Environmental Modelling & Software **25**(10): 1166-1175.

Arnot, J. A., D. Mackay, et al. (2006). "Screening Level Risk Assessment Model for Chemical Fate and Effects in the Environment." Environmental Science & Technology **40**(7): 2316-2323.

Cohen Hubal, E. A., A. Richard, et al. (2010). "Advancing Exposure Characterization for Chemical Evaluation and Risk Assessment." Journal of Toxicology and Environmental Health, Part B: Critical Reviews **13**(2): 299 - 313.

Dix, D. J., K. A. Houck, et al. (2007). "The ToxCast program for prioritizing toxicity testing of environmental ch emicals." Toxicological Sciences **95**(1): 5-12.

Egeghy, P. P., D. A. Vallero, et al. (2011). "Mandates for Exposure-Based Prioritization of Chemicals for Risk Assessment." Science of The Total Environment **in press**.

Georgopoulos, P. G. and P. J. Lioy (2006). "From a Theoretical Framework of Human Exposure and Dose Assessment to Computational System Implementation: The Modeling ENvironment for TOtal Risk Studies (MENTOR)." Journal of Toxicology and Environmental Health, Part B **9**(6): 457-483.

Hellweg, S., E. Demou, et al. (2009). "Integrating human indoor air pollutant exposure within Life Cycle Impact Assessment." Environmental Science & Technology **43**(6): 1670-1679.

Jayjock, M. A., C. F. Chaisson, et al. (2009). "Using publicly available information to create exposure and risk-based ranking of chemicals used in the workplace and consumer products." Journal of Exposure Science and Environmental Epidemiology **19**(5): 515-524.

Kephalopoulos, S. D., A. Arvanitis, et al. (2008). "GExFRAME-a Web-Based Framework for Accommodating Global Consumer Exposure Data, Scenarios and Models." Epidemiology **19**(6): S182-S183 110.1097/1001.ede.0000340054.0000300947.0000340021.

Muir, D.C.G. and P.H. Howard (2006). "Are there other persistent organic pollutants? A challenge for environmental chemists." Environmental Science & Technology **40**: 7157–7166.

NRC (2009). Science and Decisions: Advancing Risk Assessment. Washington, DC, The National Academies Press.

NRC (National Research Council) (1983). Risk Assessment in the Federal Government: Managing the Process. Washington, DC, The National Academies Press.

Rosenbaum, R., T. Bachmann, et al. (2008). "USEtox—the UNEP-SETAC toxicity model: recommended characterisation factors for human toxicity and freshwater ecotoxicity in life cycle impact assessment." The International Journal of Life Cycle Assessment **13**(7): 532-546.

Schinkel, J., N. Warrenet al. (2011). "Advanced REACH Tool (ART): calibration of the mechanistic model." Journal of Environmental Monitoring **13**(5):1374-82.

Swanson, M. B. and A. C. Socha, Eds. (1997). Chemical Ranking and Scoring: Guidelines for Relative Assessments of Chemicals. Proceedings of the Pellston Workshop on Chemical Ranking and Scoring. Pensacola, FL, SETAC Press.

Wenger, Y., L. Dingsheng, et al. (2012). "Indoor intake fraction considering surface sorption of air organic compounds for life cycle assessment." International Journal o fLife Cycle Assessment **17**(7): 919-931.

Table 1 – Model Challenge Chemicals

Chemical Name	CASRN	Chemical Name	CASRN
1,2,3–Trichlorobenzene	87-61-6	Hexabromocyclododecane	25637-99-4
2,4,5-Trichlorophenoxyacetic acid	93-76-5	Hexachlorobenzene	118-74-1
2,4-D	94-75-7	Lead	7439-92-1
8-2 fluorotelomer acid	27854-31-5	Malathion	121-75-5
Aldicarb	116-06-3	Manganese	7439-96-5
Aroclor_1254	11097-69-1	Methoxychlor	72-43-5
Aroclor_1260	11096-82-5	Methyl Mercury	22967-92-6
Arsenic	7440-38-2	Methylparaben	99-76-3
Atrazine	1912-24-9	n-Hexane	110-54-3
Benzene, 1-methoxy-4-(2-propen-1-yl)-	140-67-0	Nonylphenol	25154-52-3
Bisphenol-A	80-05-7	octaBDE	32536-52-0
Butylhydroxyanisole	8003-24-5	Parathion	56-38-2
C10-13 Chloroalkanes	85535-84-8	Pentachlorophenol	87-86-5
Cadmium	7440-43-9	pentaDBE	32534-81-9
Carbaryl	63-25-2	Perchlorate	10034-81-8 [1]
DDT	50-29-3	Permethrin	52645-53-1
decaBDE	1163-19-5	PFOS	1763-23-1
DEHP, Di(2-ethylhexyl)phthalate	117-81-7	Phenol, (1,1-dimethylethyl)-4-methoxy-	25013-16-5
Diethyl Phthalate	84-66-2	p-tert-Pentylphenol	80-46-6
Di-n-butylphtalate	84-74-2	Styrene	100-42-5
Ethane, 1,1,2,2-tetrachloro-	79-34-5	Tetrabromobisphenol A	79-94-7
Ethene, 1,1,2,2-tetrachloro- (perc)	127-18-4	Trifluralin	1582-09-8
Ethylene thiourea	96-45-7	Tris (1,3-dichloro-2-propyl) phosphate	13674-87-8
Ethylparaben	120-47-8	Tris (2-chloroethyl) phosphate	115-96-8
Formaldehyde	50-00-0	Vinclozolin	50471-44-8
gamma-Hexachlorocyclohexane	58-89-9	Vinyl Chloride	75-01-4

[1] The CASRN was incorrectly provided to the challenge participants as 10034-81-8. The correct number is 14797-73-0

Table 2 – Summary of Comparative Analysis

Type of Emission Assumption	Models	Metrics	Method	Number of Chemicals Evaluated
Actual Emissions	RAIDAR, USEtox and PRoTEGE - median estimates and 95th percentile for RAIDAR and Protégé	iR_A, C_A, exposure dose	Comparison Table	4
Actual Emissions	USEtox and PRoTEGE - median estimates	iR_A, exposure dose	Correlation	18
Actual Emissions	Binned data - RAIDAR, USEtox and PRoTEGE (median estimates and 95th percentile for RAIDAR and Protégé) with PRoTEGE semi-quantitative	iR_A, C_A, exposure dose	Correlation	16
Unit Emissions to water	RAIDAR, USEtox and EFAST	iF, C_U, LADD	Correlation /Friedman test	38
Unit Emissions to air	RAIDAR and USEtox	iF, C_U	Correlation /Friedman test	38
Unit Emissions for the sum of air and water	RAIDAR and USEtox	iF, C_U	Correlation /Friedman test	38

Table 3a – Comparison of Exposure Metrics Based on Actual Emission Data Analysis Plan

Actual Emissions Comparisons			
	RAIDAR	**USETOX**	**PRoTEGE**
Amount of Emission:	European Union (EU) – Technical Guidance Document (TGD) Actual Emissions or US EPA Pesticide emissions data for 10 chemicals	US EPA Pesticide emissions data, National Emissions Inventory (NEI) for air, Toxic Release Inventory (TRI) for air, and TRI for water	Multimedia contaminant concentrations from national databases like TRI and NEI
Mode of Entry:	EU-TGD Mode Of Entry (MOE) Estimates	US EPA Pesticide to crop application rate, NEI air, TRI air, TRI water	Scenario Specific
Aggregates Sources:	All sources (Foliage, Root, Fish, Poultry, Pork, Cow, Dairy, Milk, Egg, Inhalation, Water, Total Foods, Dust)	Sum of Water and Air - all sources (urban air, rural air, continental freshwater) – crops, root crops, meat milk and fish from freshwater and marine compartments, etc.	Ambient air, concentrations in food, environmental field studies, drinking water, and indoor air
Exposure Pathway:	Inhalation and Ingestion	Inhalation and ingestion of (urban air, rural air, continental freshwater)	Inhalation & Ingestion
Metric:	individual intake rate (iR) (μg/day) and 95th percentile based on confidence score	Population intake rate, iR (kg/year)	g/day for the population

Table 3b - – Comparison of Exposure Metrics Based on Unit Emission Data

Unit Emissions Comparisons		
	RAIDAR	
Amount of Emission:	Unit	Unit
Mode of Entry:	Water	Air
Source:	Foliage, Root, Fish,TerrHerb, TerrCarn, Poultry, Pork, Cow, Dairy, Milk, Egg, Inhalation, Water, Total Foods	Foliage, Root, Fish,TerrHerb, TerrCarn, Poultry, Pork, Cow, Dairy, Milk, Egg, Inhalation, Water, Total Foods
Pathway:	Inhalation and Ingestion	Inhalation and Ingestion
Metric:	Individual, *iF* (kg/kg), iR_U (μg/day), C_U (μg/kg)	Individual, *iF* (kg/kg), $\underline{iR_U}$ (μg/day), C_U (μg/kg)
	USETOX	
Amount of Emission:	Unit	Unit
Mode of Entry:	Freshwater	Continental rural air
Source:	Freshwater - crops, root crops, meat milk and fish from freshwater and marine compartments, etc.	Rural air - crops, root crops, meat milk and fish from freshwater and marine compartments, etc
Pathway:	Inhalation and Ingestion	Inhalation and Ingestion
Metric:	iF_{POP} (kg/kg)	iF_{POP} (kg/kg)
	EFAST	
	Overall human exposure estimates	
Amount of Emission:	Unit (1 kg over 20 days) for all chemicals	
Mode of Entry:	Landfill infiltration to groundwater, industrial fugitive gas and stack releases to air, surface water releases	
Source:	Drinking water, fish ingestion, inhalation and groundwater ingestion	
Pathway:	Inhalation and Ingestion	
Metric:	mg/kg/day	

Table 4 – Categorical Results from GExFRAME

Category	Chemical Description	Human Exposure Assessment Process
1	Previously used and currently not found in human exposure media	No human exposure possible
2	Previously used and currently found in human exposure media	Exposure assessment for chemicals found in near field exposure media
3	Present in human exposure media during industrial use	Worker exposure only during the manufacturing process for both raw material and products
4	Found in food and/or drinking water	Exposure assessment for chemicals present in food and/or drinking water
5	Found in consumer use products	Exposure assessment for chemicals present in consumer products resulting in their presence in near field exposure media
6	A pesticide currently in use	Exposure assessment for pesticides as done by OPP (Office of Pesticide Programs)

Table 5 – Comparative Ranks of Chemicals with US Actual Emission Estimates

Chemical Name	RAIDAR	RAIDAR	USETOX	PROTÉGÉ	PROTÉGÉ
	Intake Fraction	95th	Intake Fraction	Median	95th
	ug/day	ug/day	kg·/yr-1	g/day	g/day
2,4-D	2	2	1	1	2
Aldicarb	3	3	2	2	1
Methoxychlor	1	1	3	3	3
Parathion	4	4	4	4	4

Table 6 – Friedman Test Results for Unit Emission Model Rankings

Measurement	Models	Metrics	Friedman chi-squared	Degrees of Freedom	p-value
Unit Emissions to water	RAIDAR, USEtox and EFAST	iF, C_U, LADD	0.0771	df = 3,	0.9944
Unit Emissions to air	RAIDAR and USEtox	iF, C_U	0.5	df = 2	0.7788
Unit Emissions for the sum of air and water	RAIDAR and USEtox	iF, C_U	0.9313	df = 2	0.6277

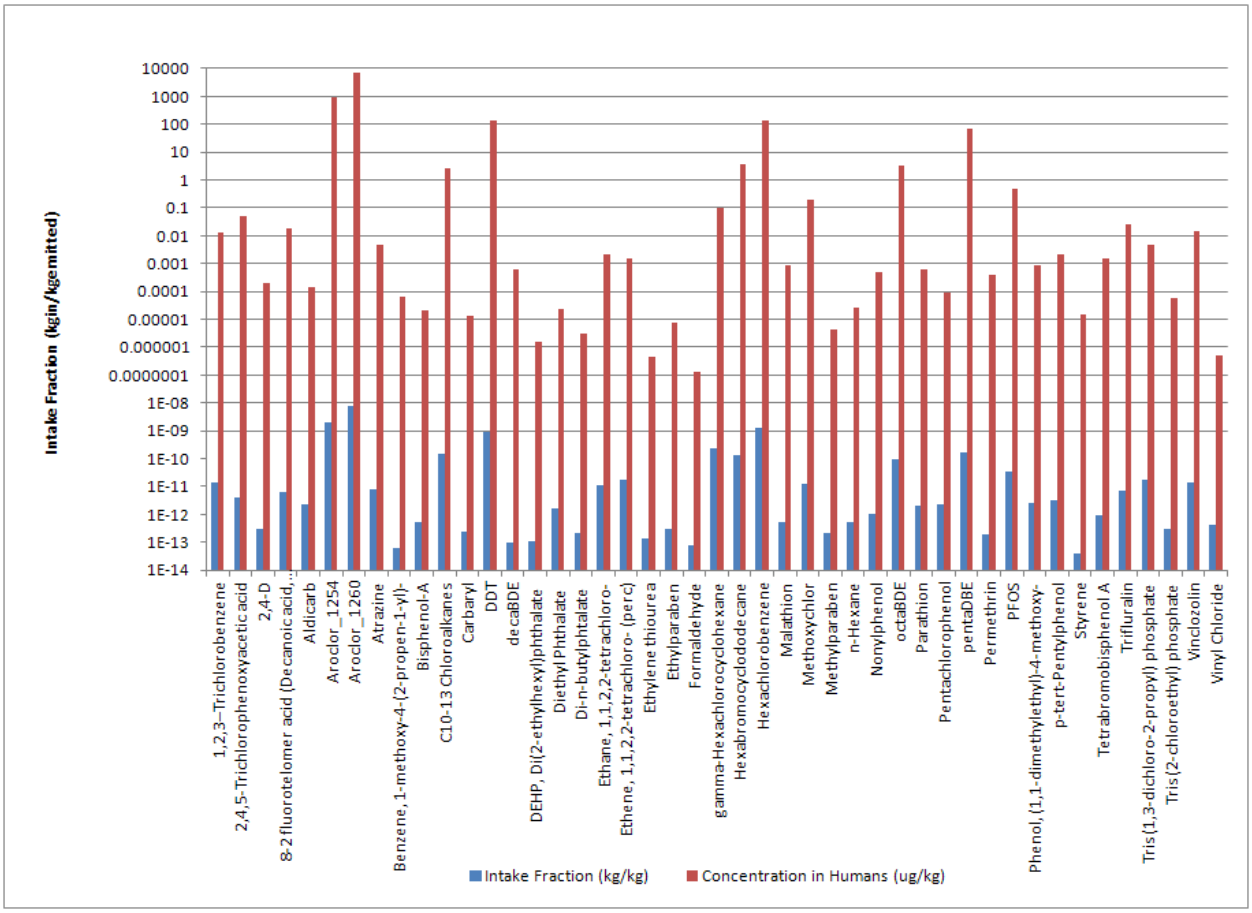


Figure 1 – Comparison between estimates produced by RAIDAR exposure metrics

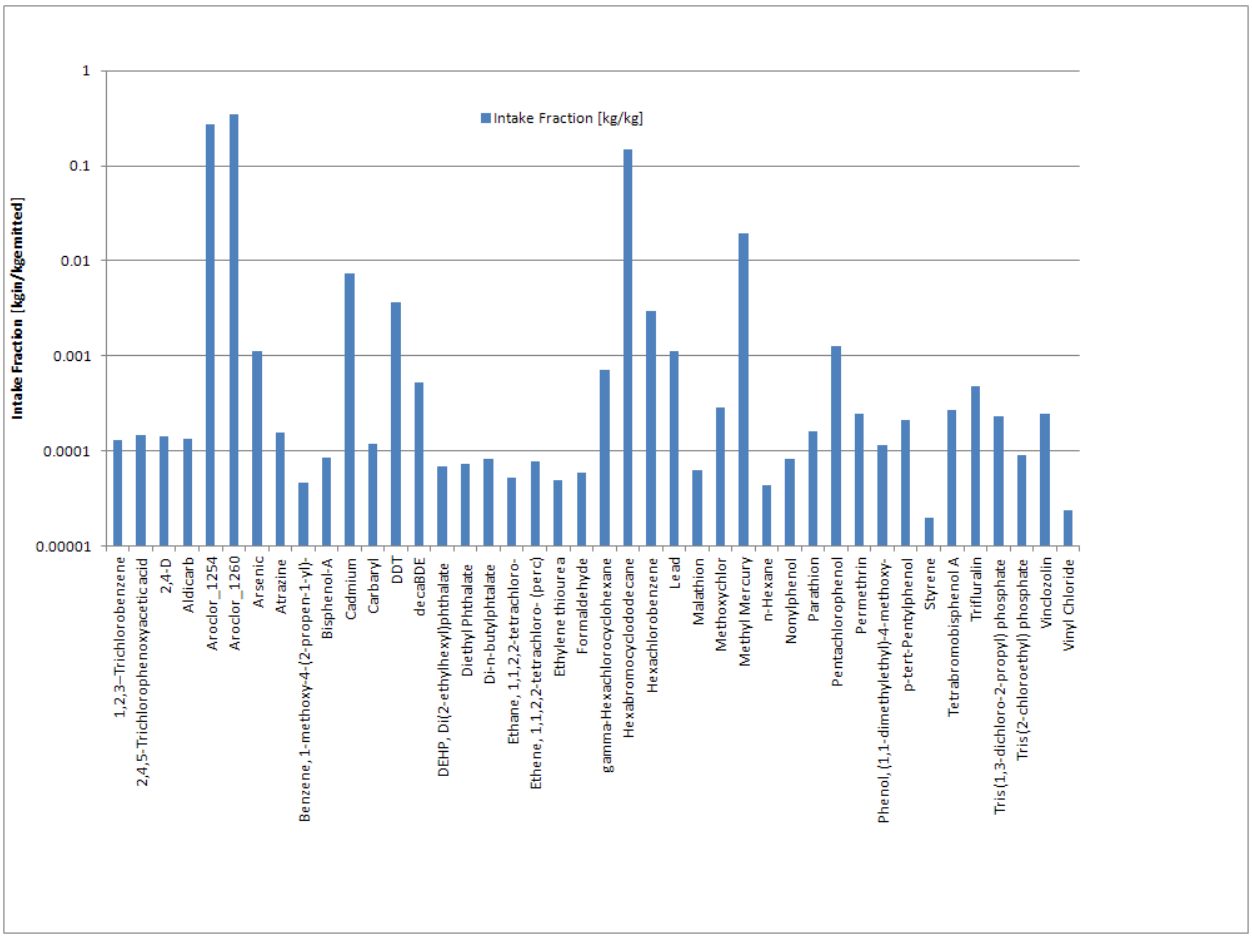


Figure 2 – Summary of Rankings Produced by USEtox

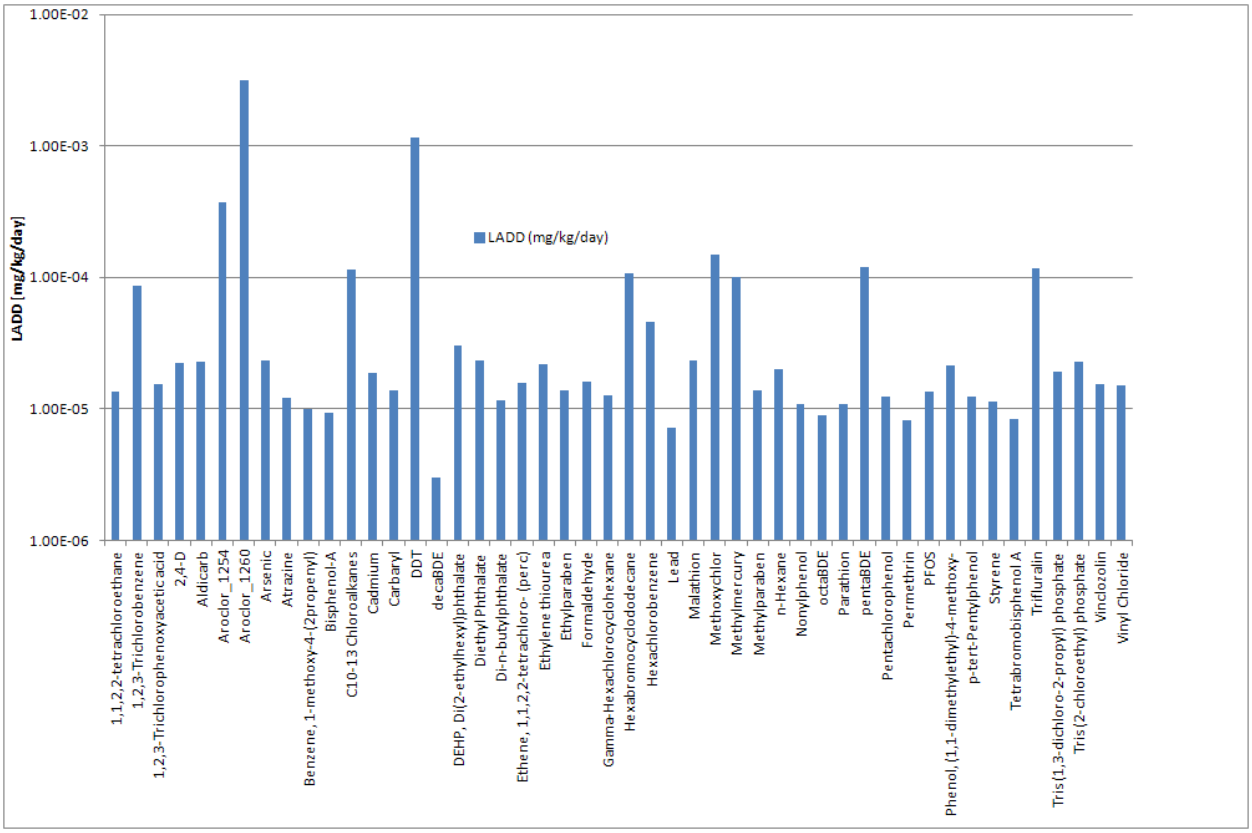


Figure 3 - Summary of Rankings Produced by EFAST2

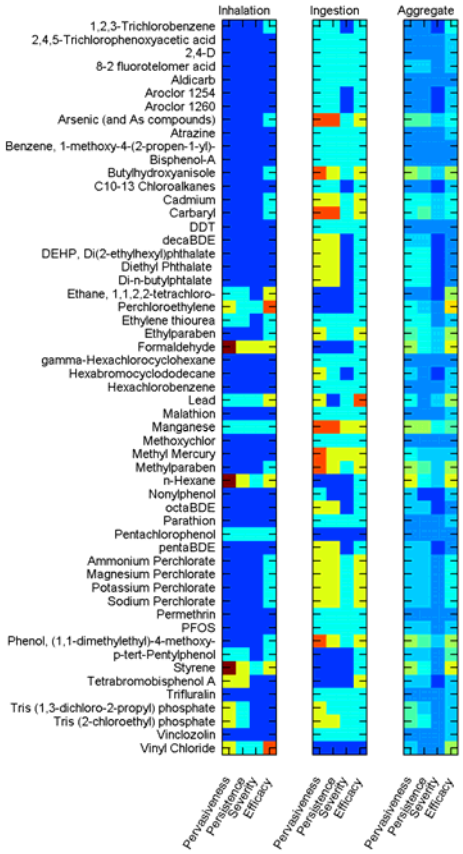


Figure 4 - Summary of "Tier 1" estimates of semi-quantitative metrics of exposure (pervasiveness, persistence, severity, and efficacy) for the 55 chemicals

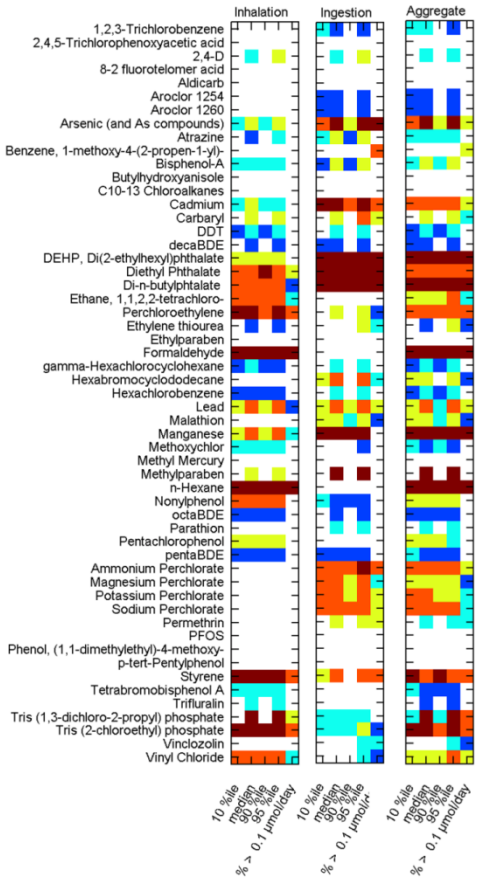
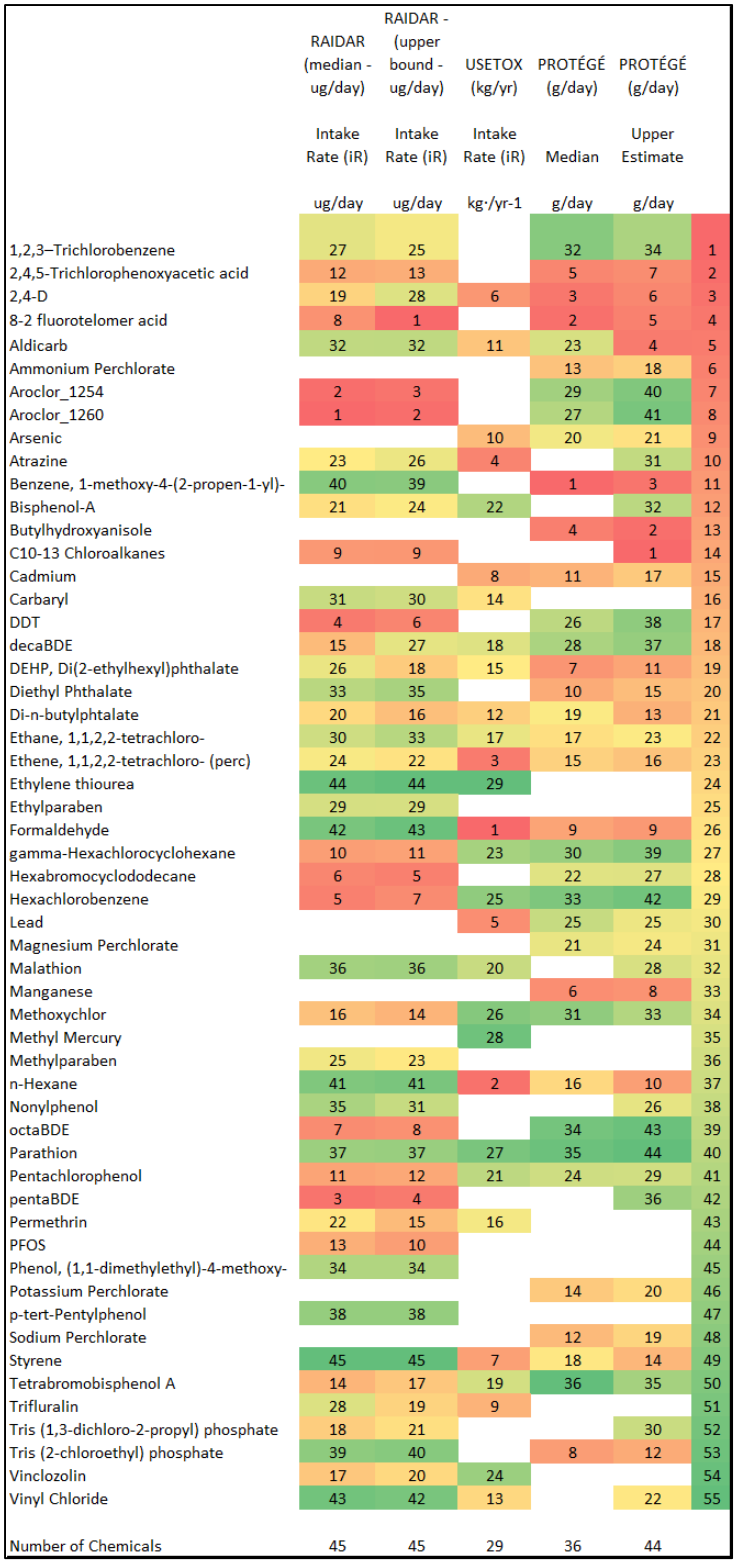
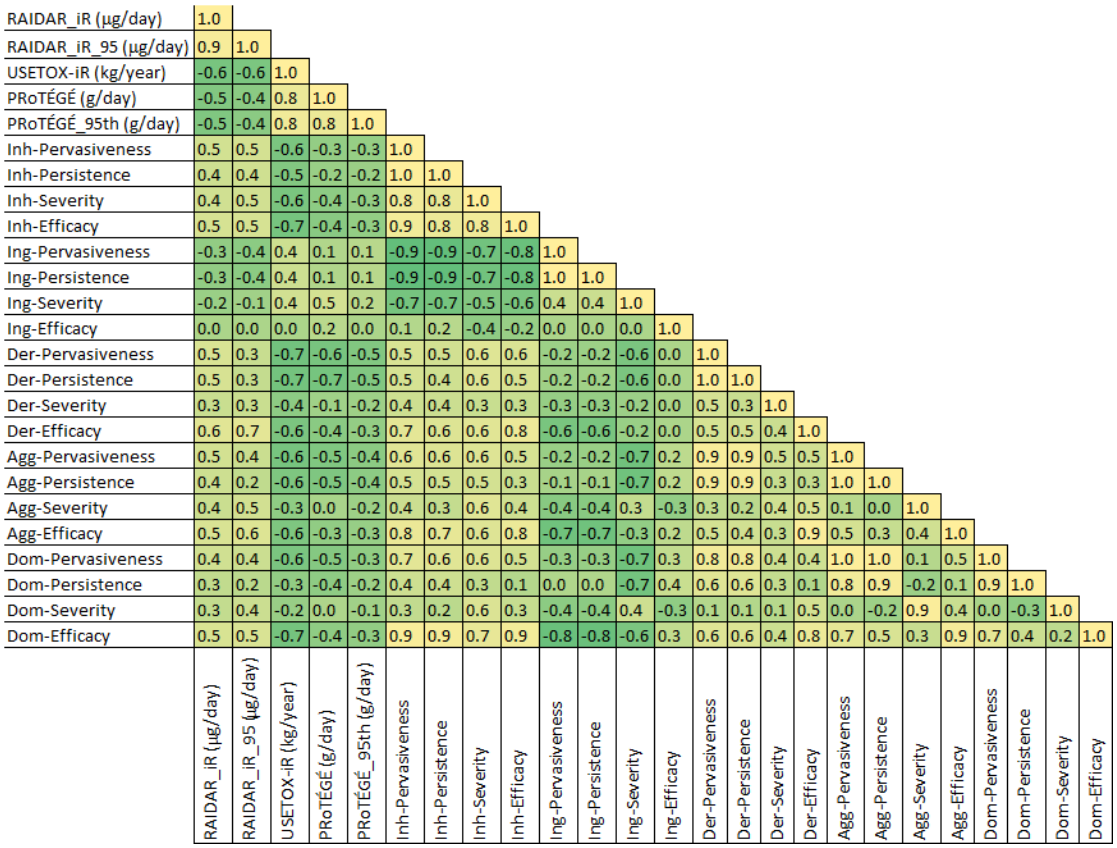


Figure 5 - Summary of "Tier 2" estimates of quantitative metrics of exposure of the general population of the US for the 55 chemicals. Rankings are color coded based on values from 1 to 5 as shown in the legend and white (empty) cells indicate no data (ND) were available.



	RAIDAR (median - ug/day) Intake Rate (iR) ug/day	RAIDAR - (upper bound - ug/day) Intake Rate (iR) ug/day	USETOX (kg/yr) Intake Rate (iR) kg·/yr-1	PROTÉGÉ (g/day) Median g/day	PROTÉGÉ (g/day) Upper Estimate g/day	
1,2,3–Trichlorobenzene	27	25		32	34	1
2,4,5-Trichlorophenoxyacetic acid	12	13		5	7	2
2,4-D	19	28	6	3	6	3
8-2 fluorotelomer acid	8	1		2	5	4
Aldicarb	32	32	11	23	4	5
Ammonium Perchlorate				13	18	6
Aroclor_1254	2	3		29	40	7
Aroclor_1260	1	2		27	41	8
Arsenic			10	20	21	9
Atrazine	23	26	4		31	10
Benzene, 1-methoxy-4-(2-propen-1-yl)-	40	39		1	3	11
Bisphenol-A	21	24	22		32	12
Butylhydroxyanisole				4	2	13
C10-13 Chloroalkanes	9	9			1	14
Cadmium			8	11	17	15
Carbaryl	31	30	14			16
DDT	4	6		26	38	17
decaBDE	15	27	18	28	37	18
DEHP, Di(2-ethylhexyl)phthalate	26	18	15	7	11	19
Diethyl Phthalate	33	35		10	15	20
Di-n-butylphtalate	20	16	12	19	13	21
Ethane, 1,1,2,2-tetrachloro-	30	33	17	17	23	22
Ethene, 1,1,2,2-tetrachloro- (perc)	24	22	3	15	16	23
Ethylene thiourea	44	44	29			24
Ethylparaben	29	29				25
Formaldehyde	42	43	1	9	9	26
gamma-Hexachlorocyclohexane	10	11	23	30	39	27
Hexabromocyclododecane	6	5		22	27	28
Hexachlorobenzene	5	7	25	33	42	29
Lead			5	25	25	30
Magnesium Perchlorate				21	24	31
Malathion	36	36	20		28	32
Manganese				6	8	33
Methoxychlor	16	14	26	31	33	34
Methyl Mercury			28			35
Methylparaben	25	23				36
n-Hexane	41	41	2	16	10	37
Nonylphenol	35	31			26	38
octaBDE	7	8		34	43	39
Parathion	37	37	27	35	44	40
Pentachlorophenol	11	12	21	24	29	41
pentaBDE	3	4			36	42
Permethrin	22	15	16			43
PFOS	13	10				44
Phenol, (1,1-dimethylethyl)-4-methoxy-	34	34				45
Potassium Perchlorate				14	20	46
p-tert-Pentylphenol	38	38				47
Sodium Perchlorate				12	19	48
Styrene	45	45	7	18	14	49
Tetrabromobisphenol A	14	17	19	36	35	50
Trifluralin	28	19	9			51
Tris (1,3-dichloro-2-propyl) phosphate	18	21			30	52
Tris (2-chloroethyl) phosphate	39	40		8	12	53
Vinclozolin	17	20	24			54
Vinyl Chloride	43	42	13		22	55
Number of Chemicals	45	45	29	36	44	

Figure 6 – Summary Comparison of Rankings for Each Chemical Based on an Actual Emission Rate – Number in each column represents the rank of the chemical among the set of chemicals ranked by each model and/or metric. Shading varies from high (red) to low (green) exposure based on the estimates provided.



	RAIDAR_iR (µg/day)	RAIDAR_iR_95 (µg/day)	USETOX-iR (kg/year)	PRoTÉGÉ (g/day)	PRoTÉGÉ_95th (g/day)	Inh-Pervasiveness	Inh-Persistence	Inh-Severity	Inh-Efficacy	Ing-Pervasiveness	Ing-Persistence	Ing-Severity	Ing-Efficacy	Der-Pervasiveness	Der-Persistence	Der-Severity	Der-Efficacy	Agg-Pervasiveness	Agg-Persistence	Agg-Severity	Agg-Efficacy	Dom-Pervasiveness	Dom-Persistence	Dom-Severity	Dom-Efficacy
RAIDAR_iR (µg/day)	1.0																								
RAIDAR_iR_95 (µg/day)	0.9	1.0																							
USETOX-iR (kg/year)	-0.6	-0.6	1.0																						
PRoTÉGÉ (g/day)	-0.5	-0.4	0.8	1.0																					
PRoTÉGÉ_95th (g/day)	-0.5	-0.4	0.8	0.8	1.0																				
Inh-Pervasiveness	0.5	0.5	-0.6	-0.3	-0.3	1.0																			
Inh-Persistence	0.4	0.4	-0.5	-0.2	-0.2	1.0	1.0																		
Inh-Severity	0.4	0.5	-0.6	-0.4	-0.3	0.8	0.8	1.0																	
Inh-Efficacy	0.5	0.5	-0.7	-0.4	-0.3	0.9	0.8	0.8	1.0																
Ing-Pervasiveness	-0.3	-0.4	0.4	0.1	0.1	-0.9	-0.9	-0.7	-0.8	1.0															
Ing-Persistence	-0.3	-0.4	0.4	0.1	0.1	-0.9	-0.9	-0.7	-0.8	1.0	1.0														
Ing-Severity	-0.2	-0.1	0.4	0.5	0.2	-0.7	-0.7	-0.5	-0.6	0.4	0.4	1.0													
Ing-Efficacy	0.0	0.0	0.0	0.2	0.0	0.1	0.2	-0.4	-0.2	0.0	0.0	0.0	1.0												
Der-Pervasiveness	0.5	0.3	-0.7	-0.6	-0.5	0.5	0.5	0.6	0.6	-0.2	-0.2	-0.6	0.0	1.0											
Der-Persistence	0.5	0.3	-0.7	-0.7	-0.5	0.5	0.4	0.6	0.5	-0.2	-0.2	-0.6	0.0	1.0	1.0										
Der-Severity	0.3	0.3	-0.4	-0.1	-0.2	0.4	0.4	0.3	0.3	-0.3	-0.3	-0.2	0.0	0.5	0.3	1.0									
Der-Efficacy	0.6	0.7	-0.6	-0.4	-0.3	0.7	0.6	0.6	0.8	-0.6	-0.6	-0.2	0.0	0.5	0.5	0.4	1.0								
Agg-Pervasiveness	0.5	0.4	-0.6	-0.5	-0.4	0.6	0.6	0.6	0.5	-0.2	-0.2	-0.7	0.2	0.9	0.9	0.5	0.5	1.0							
Agg-Persistence	0.4	0.2	-0.6	-0.5	-0.4	0.5	0.5	0.5	0.3	-0.1	-0.1	-0.7	0.2	0.9	0.9	0.3	0.3	1.0	1.0						
Agg-Severity	0.4	0.5	-0.3	0.0	-0.2	0.4	0.3	0.6	0.4	-0.4	-0.4	0.3	-0.3	0.3	0.2	0.4	0.5	0.1	0.0	1.0					
Agg-Efficacy	0.5	0.6	-0.6	-0.3	-0.3	0.8	0.7	0.6	0.8	-0.7	-0.7	-0.3	0.2	0.5	0.4	0.3	0.9	0.5	0.3	0.4	1.0				
Dom-Pervasiveness	0.4	0.4	-0.6	-0.5	-0.3	0.7	0.6	0.6	0.5	-0.3	-0.3	-0.7	0.3	0.8	0.8	0.4	0.4	1.0	1.0	0.1	0.5	1.0			
Dom-Persistence	0.3	0.2	-0.3	-0.4	-0.2	0.4	0.4	0.3	0.1	0.0	0.0	-0.7	0.4	0.6	0.6	0.3	0.1	0.8	0.9	-0.2	0.1	0.9	1.0		
Dom-Severity	0.3	0.4	-0.2	0.0	-0.1	0.3	0.2	0.6	0.3	-0.4	-0.4	0.4	-0.3	0.1	0.1	0.1	0.5	0.0	-0.2	0.9	0.4	0.0	-0.3	1.0	
Dom-Efficacy	0.5	0.5	-0.7	-0.4	-0.3	0.9	0.9	0.7	0.9	-0.8	-0.8	-0.6	0.3	0.6	0.6	0.4	0.8	0.7	0.5	0.3	0.9	0.7	0.4	0.2	1.0

Figure 7 – Correlation of Binned Chemical Rankings and Qualitative Metric – Spearman Coefficient between each model and metric used to estimate exposure for the same set of chemicals. Shading indicates degree of correlation from high (yellow) to low (green). Abbreviations: Inhalation (Inh), Ingestion (Ing), Dermal Absorption (Der), Aggregate Pathways (Agg), Dominant Pathway (Dom)

RAIDAR_IF	1.0			
RAIDAR_CU	0.9	1.0		
USETOX_IF	0.7	0.7	1.0	
EFAST_LADD	0.3	0.3	0.3	1.0
	RAIDAR_IF	RAIDAR_CU	USETOX_IF	EFAST_LADD

Figure 8 – Correlation of Unit Emission to Water Rankings - Spearman Coefficient between each model and metric used to estimate exposure for the same set of chemicals. Shading indicates degree of correlation from high (yellow) to low (green)

RAIDAR_IF	1.0		
RAIDAR_CU	0.8	1.0	
USETOX_IF	0.9	0.7	1.0
	RAIDAR_IF	RAIDAR_CU	USETOX_IF

Figure 9 – Correlation of Unit Emission to Air Rankings - Spearman Coefficient between each model and metric used to estimate exposure for the same set of chemicals. Shading indicates degree of correlation from high (yellow) to low (green)