Proceedings of the International Symposium on Sustainable Systems and Technologies, v3 (2015)

Ranking the Relative Importance of Toxicological Observations Based on Subject Matter Expertise

Kurt A. Gust US Army Engineer Research & Dev. Center, Kurt.A.Gust@usace.army.mil
Michael L. Mayo US Army Engineer Research & Dev. Center, Michael.L.Mayo@usace.army.mil
Zachary A. Collier US Army Engineer Research & Dev. Center, Zachary.A.Collier@usace.army.mil

Abstract. Toxicological investigations provide key observations for establishing potential impacts of chemicals and other materials in environmental exposures. The specific evaluation of chemical toxicity can be drawn from numerous lines of evidence, including, but, not limited to: lethality, sub-lethal impacts, toxicokinetics, critical tissue residues, molecular toxicology, reproductive impacts, mutation, carcinogenesis, decreased genetic diversity in populations and decreased species diversity. Despite this richness of potential observations, many predictive toxicological models, such as those included in life cycle impact assessment (LCIA), generally rely on one or a few ecotoxicity measures including median lethal concentrations (LC_{50}) and no observed effects concentrations (NOECs). By excluding all but these composite effects metrics, such models often fail to accurately and/or robustly capture potential adverse outcomes that may negatively affect individual health, biological fitness and/or population sustainability. Although each has the potential to add value to chemical risk assessment, the relative importance of the various toxicological observations available for characterizing and assessing environmental impact has not been established. Therefore, we sought to establish the importance of multiple ecotoxicological lines of evidence for assessing environmental impacts. "Importance" in this context is a subjective valuation based on perceived utility within an assessment, therefore we developed and administered a survey to query subject matter experts (SMEs) in the fields of toxicology, ecotoxicology and risk assessment on the importance of a broad array of ecotoxicological measurements. This scoping study focused on SME opinions from Department of Defense scientists and risk assessors where 61 surveys were distributed and 21 returned and analyzed. Overall, observations from chronic exposures and reproductive endpoints were ranked the highest among SMEs while lines of evidence from *in vivo* studies outranked both *in vitro* and *in silico* observations.

Proceedings of the International Symposium on Sustainable Systems and Technologies (ISSN 2329-9169) is published annually by the Sustainable Conoscente Network. Jun-Ki Choi and Annick Anctil, co-editors 2015. ISSSTNetwork@gmail.com.

Cite as: Ranking the Relative Importance of Toxicological Observations Based on Subject Matter Expertise. *Proc. ISSST*, Gust K.A., Mayo M.L., Collier Z.A. http://dx.doi.org/10.6084/m9.figshare.1517802 (2015)

Non-adverse effects, regardless of acute or chronic exposure, ranked the lowest in importance. Curiously, SME opinions were highly variable across nearly all lines of toxicological information reflected by very broad score distributions and high variance. Survey results also indicated that respondents level of experience (<10, 10-20, and >20 years within the profession) had no significant effect on importance scoring for the majority of toxicological categories with the exception of reproduction and adverse effect parameters. We discuss the immediate value and context of these observations as well as describe how constructing these distributions can allow for emulation of a larger virtual expert population through stochastic sampling. Finally, we discuss how future virtual expert distributions might be used as a mechanism to weight multiple lines of evidence into an aggregate LCIA ecotoxicity impact factor.

Introduction. Within the toxicological and ecotoxicological sciences, a variety of methods and assays have been developed to help characterize the hazard and risk of chemicals. Methods range from highly standardized tests that are used by regulatory agencies such as the US Environmental Protection Agency (EPA) to assess environmental quality/compliance, to highly specialized and unique methods used to assess research-specific questions. There are a multitude of potential lines of evidence that may be available for understanding the toxicity and toxicology of a given compound (Owens, 1998). Given the advent of the adverse outcome pathway (AOP) concept, the means to consolidate multiple lines of evidence into singular toxicological contexts are continuously evolving. The AOP represents a conceptual framework linking a molecular initiating event (MIE) to an adverse outcome of regulatory concern (e.g., mortality, impaired reproduction, and/or other effects that negatively impact biological fitness) through a series of key events occurring at increasing levels of biological complexity (Ankley et al., 2010). Given that each key event may be tested using a number of different methods/ assays, prioritizing the relative importance of each in a resource-constrained environment (money, time, manpower, etc.) becomes a pressing issue for the scientific and regulatory community.

Therefore, a need exists for a method to determine the relative importance of results across these methods/assays for use in establishing chemical hazard and ultimately to estimate risk. Given that the exercise of ascribing *value* to scientific findings is an inherently subjective task, and that real decision processes are based on *both* empirical data and value judgments, we must turn to the realm of subjective expert judgment (Cooke & Goossens, 2008; Owens, 1998). However, since the decision problem is uncertain, experts will not all agree, and therefore a census-type approach can be taken to determine an overall distribution of subject matter opinion rather than aggregating these data into a single point-estimate (Cooke & Goossens, 2008).

Examples of stochastic approaches for comparing the relative importance of multiple indicators have been conducted by Prado-Lopez et al. (2013, 2014) across multiple life cycle impact categories (e.g., ecotoxicity vs. global warming potential vs. eutrophication). However, it is unclear how more granular components within a specific impact category (in this case, ecotoxicity) rank in terms of their relative importance (Owens, 1998).

Goals. The purpose of the present study was to determine the subjective importance of each line of evidence within an overall chemical-effects assessment. The approach to meet this goal was to develop and administer a subject matter expert (SME) survey to gather professional judgments of toxicologists, ecotoxicologists and risk assessors on the relative importance of six major classes of biological information, as they relate to ecotoxicological assessment. Based on these results, probability distributions were constructed, which may be utilized as a “virtual” population of subject matter expertise.

In particular, we sought to express the relative importance of various lines of toxicological evidence through relative weights elicited from SMEs. These weights represent the relative importance of multiple ecotoxicological indicators, and are based on the perceived value that SMEs place on those pieces of information. The definition, mathematical interpretation, and elicitation of weights in the context of value modeling have been rigorously covered elsewhere (for details, see Belton & Stewart, 2002). However, the main purpose of using subject matter expertise in this way is to use these subjective weights as scientific data themselves (Cooke & Goossens, 2008).

Investigative Method. A SME survey was developed to gather the opinions of experts in ecotoxicology, toxicology and/or risk assessment for key types of ecotoxicological information. Six major classes of biological information were identified to be pervasive in the toxicological and ecotoxicological effects assessment literature: (1) lethality, (2) sub-lethal effects, (3) absorption, distribution, metabolism and elimination (ADME), (4) systems biology, (5) population-level effects and (6) community-level effects (Figure 1). Additionally, each of these classes of biological information may be derived from three distinct lines of inquiry: (1) *in vivo*, (2) *in vitro* and (3) *in silico* observations. The SME survey was developed to evaluate the perceived value of the six classes of biological information in addition to the three lines of inquiry. See the Supplementary Information for details of the survey, including descriptions of the classes of biological information.

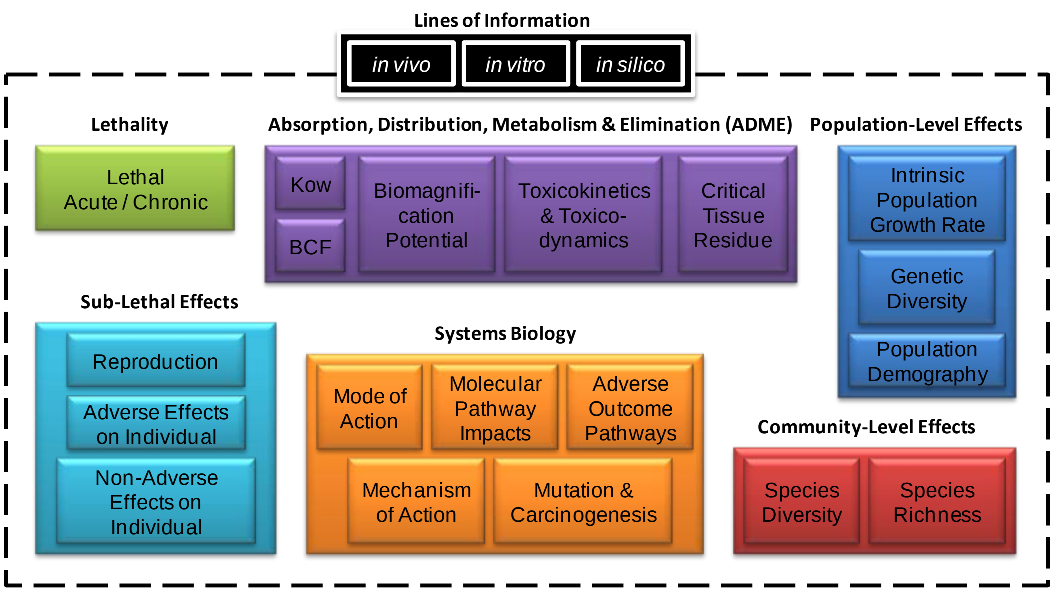

Figure 1. Classes of Ecotoxicological Information. Lines of commonly considered ecotoxicological information divided into six general categories.

Participant Scoring Instructions. The survey was focused on only ranking the benefits associated from gaining the particular information contained within a metric of interest, allowing the SMEs to score a particular line of information independent of the cost, time and effort required to generate a particular type of information. Therefore, the scores should represent only the benefit of the particular type of information to ecotoxicological understanding irrespective of cost or technical challenge (which may have skewed the SMEs assessment of a metric's utility). This pure benefit score provides researches an unbiased assessment of the value of specific observation types upon which they can subsequently apply their own value versus cost analysis for application within risk assessment. Finally, all metrics were scored

independently. Specifically, the SMEs were informed to score the survey in a manner where their score for one type of information was not dependent on the scoring of any other information type across the survey. This value scoring framework, consistent with a “Value Focused Thinking” approach (Keeney, 2009), requires the participants to focus on the reasons why one piece of information may be more beneficial than another, as opposed to remaining in entrenched positions about specific alternatives such as favored tests or metrics (Collier et al., 2014).
Subject to the above requirements, participants were asked to score the six classes and respective subclasses of biological information using a 0-100 scale, wherein a score of “0” represented the lowest importance and “100” represented the highest (Supplemental Figure S1). Additionally, participants were asked to rank the 3 lines of biological information (i.e., *in vivo*, *in vitro*, and *in silico*), wherein “1” represents the highest ranking and “3” the lowest.

Survey Administration. Two pools of SMEs with expertise spanning toxicology, ecotoxicology, environmental risk assessment and human health risk assessment were queried to generate the results and distributions of SME opinion. Given that this work directly supports a research effort seeking to expand life cycle assessment capabilities for emerging military materials, we focused our survey administration upon DOD scientists with direct experience in assessing the effects and risk of military materials. The first pool of subject matter experts surveyed consisted of federal employees with relevant expertise within the US Army, from the Environmental Laboratory within the US Army Engineer Research and Development Center (ERDC). The second pool of SMEs queried were members of the Tri-Services Toxicology Consortium, which brings together toxicological experts from the US Army, US Navy, and US Air Force. Given the common thread among SMEs of employment within military organizations, the survey results best describe the professional opinions of SMEs in support of military science, technology and risk assessment. However, the results of this work can additionally be considered a subset of the overarching disciplines of toxicology, ecotoxicology and environmental / human risk assessment and thus represent a specifically focused, yet relevant sampling of the broader community of practice.

Distributions of Subject Matter Expert Opinions. Individual scores, labeled here by S, that were reported for each of the 6 primary classes of biological information and respective subclasses, as explained above, ranged in value between 0 and 100. A discrete probability distribution for the scores, $p_i(S)$, was estimated for each subclass i, using the Freedman-Diaconis algorithm (Freedman and Diaconis, 1981) to estimate the number of discretized intervals in support of each distribution. Data from each subclass were considered and analyzed independently of all other subclass data. Expectation values for the scores of each subclass, $\langle S \rangle_i$, were calculated according to the following equation:

$$\langle S \rangle_i = \sum_S S \times p_i(S). \tag{1}$$

The variance in scores associated with each category, σ_i^2, was calculated from these discrete probability distribution according to the equation:

$$\sigma_i^2 = \langle S^2 \rangle_i - \langle S \rangle_i^{\,2}, \tag{2}$$

wherein the second moment of the distribution, $\left\langle S^2 \right\rangle_i$, was calculated similarly to the first moment (Eq. (1)). Mean scores obtained from Eq. (1) were used to rank-order the subclasses of biological information, which included both acute and chronic exposure durations. The standard deviation of these data is the square root of the variance, $\sqrt{\sigma_i^2}$, and was used as a visual aid in addition to the mean scores to assess the statistical spread in value of the scores.

Results. Surveys were distributed to 34 SMEs in the first expert pool of which 15 surveys were completed and returned. In the second pool, 27 SME surveys were distributed and 6 surveys were completed and returned. The total number of surveys collected was therefore equal to 21, with a response rate of approximately 34%.

Ranking Classes of Biological Information. The simple ranking of SME opinions on the importance of the various lines of ecotoxicological information indicated a number of important trends. For example, information from chronic exposure assays were broadly considered to be the most important data, wherein 8 of the top 10 categories represented observations from chronic duration tests (Table 1). Out of these 8 chronic subclass scores, 4 reflected observations from assays that examined effects on reproduction. Therefore, chronic exposure assays investigating the reproduction endpoint were considered to be the most important toxicological information for conducting an effects assessment. In addition to reproduction, the ADME category appeared twice in the top 10, which provided the top-scoring subclass (chronic information for critical tissue residue). Finally, the systems biology and lethality classes made showings in the top 10, respectively including mutations/carcinogens and NOEC (chronic). SMEs considered non-adverse effects on individuals to be the least important category, regardless of acute or chronic exposure, for effects assessment. Moreover, an overview of the top 10 rated acute observations demonstrated that subject matter experts found ADME (critical tissue residues), reproduction (NOEC and ED_{50}), and median lethal dose (LD_{50}) to be the most important short duration assays for assessing biological effects. Finally, SMEs broadly considered *in vivo* observations to be the most important for toxicological assessment.

Table 1. Trends in the Opinions of Subject Matter Expert Responses

Class	Metric	Acute/Chronic	Mean	Standard Deviation
Top 10 Highest Rated Categories				
ADME	Critical tissue residue	Chronic	73.3553	14.1715
Reproduction	ED50	Chronic	72.8571	16.6599
Reproduction	NOEC	Chronic	69.9405	19.6338
Systems Biology	Mutation/Carcinogen	Chronic	69.4	13.9298
ADME	Critical tissue residue	Acute	68.8421	17.3426
Reproduction	Benchmark Dose	Chronic	68.55	18.8347
Reproduction	Species Sensitivity Distribution	Chronic	68.0952	19.4248
Reproduction	NOEC	Acute	67.2976	22.7266
Lethality	NOEC	Chronic	67.2222	21.4879
Bottom 10 Lowest Rated Categories				
Non-Adverse Effects	ED50	Acute	38.5238	19.4975
Non-Adverse Effects	NOEC	Acute	39.3036	23.5124
Non-Adverse Effects	Benchmark Dose	Acute	40	22.7684
Non-Adverse Effects	Benchmark Dose	Chronic	41.2	23.3615
Non-Adverse Effects	ED50	Chronic	41.3095	21.0953
Non-Adverse Effects	NOEC	Chronic	42.6429	26.7605
Non-Adverse Effects	Species Sensitivity Distribution	Acute	44.0179	21.4746
Non-Adverse Effects	Species Sensitivity Distribution	Chronic	45.1964	21.7956
Lethality	Benchmark Dose	Acute	49.5	25.5881
Population Level Effects	Population Demographics		53.3929	24.6196
Top 10 Highest Rated Acute Tests				
ADME	Critical tissue residue	Acute	68.8421	17.3426
Reproduction	NOEC	Acute	67.2976	22.7266
Reproduction	ED50	Acute	65.1786	23.5227
Lethality	LD50	Acute	63.244	23.7231
Adverse Effects	Species Sensitivity Distribution	Acute	63.0357	22.2692
Lethality	NOEC	Acute	62.6984	28.1271
Systems Biology	Mutation/Carcinogen	Acute	62.4	17.7268
Adverse Effects	NOEC	Acute	61.8571	23.2221
Lethality	Species Sensitivity Distribution	Acute	61.6032	23.716
Systems Biology	Adverse Outcome Pathways (AOP)	Acute	61.3	21.506

Distribution of scores within each subclass. Figure 2 illustrates the distribution of scores based on SME elicitation. Scores throughout all category subclasses are spread broadly across the interval, although some subclasses, such as acute and chronic information relating to critical tissue residue (ADME category), exhibit a smaller variance. Despite the presence of several outliers, the non-adverse effects category is rated substantially lower than all others, although the variance is similar to that observed from the other categories.

An interesting result gained from the SME opinion survey was the high level of variance observed in category scores (Figure 2). In general, the score distributions for the various categories exhibited broad ranges where for most categories, scores ranging from >20 to 100 were observed.

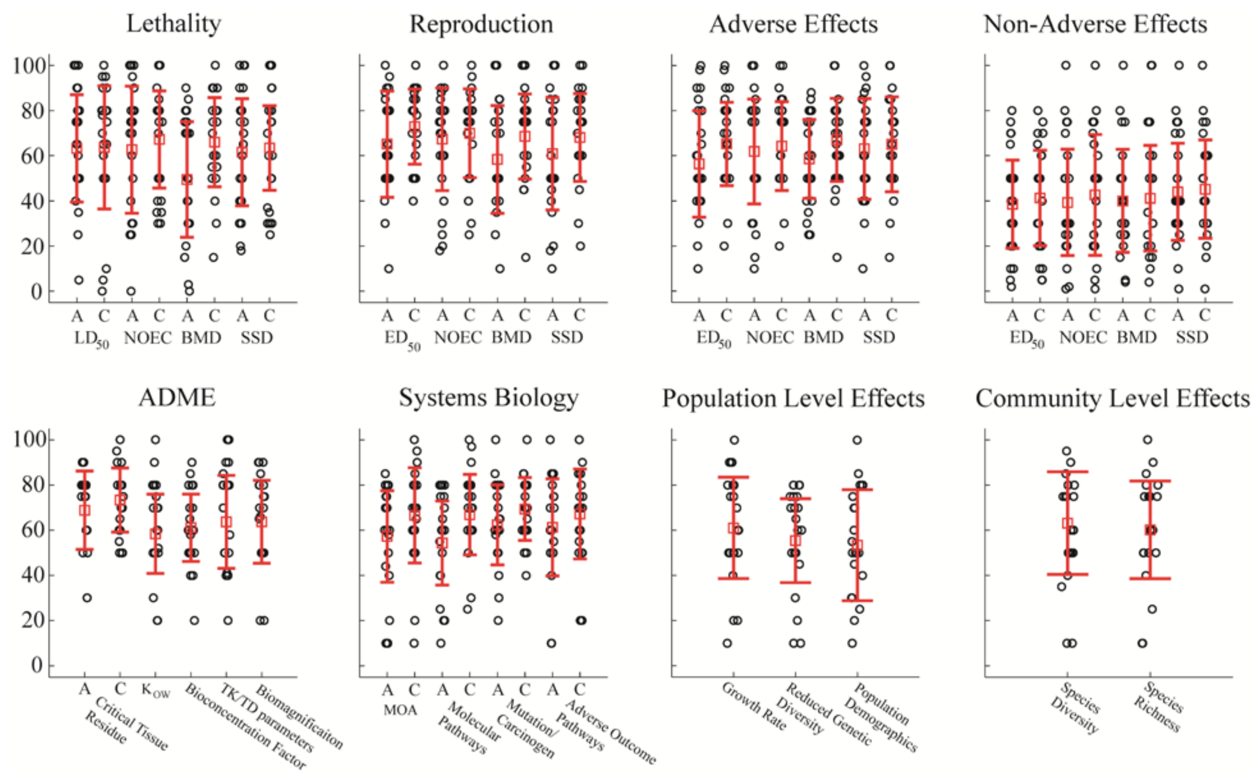

Figure 2. Subject matter expert opinion score distributions. Distributions of survey results, wherein each (black) circle represents a reported score. The red bars represent one standard deviation above and below the mean score (red squares). Here, A stands for Acute, and C stands for Chronic.

Experience level affects reproduction and adverse effect scores. We addressed whether other data obtained in the survey would confound the mean responses. In the survey, respondents were asked to provide their years of experience in their subject matter areas. Nearly all experts queried claimed expertise in toxicology or a closely related field (e.g., ecotoxicology or toxicogenomics). We carried out a one-way ANOVA to determine whether expert experience level significantly affected the score distributions. We grouped respondents into 3 experience classes: (i) early career experts, with 0-9 years; (ii) mid-career experts, with 10-19 years; and (ii) senior-level experts with >20 years experience. Interestingly, we found that mean scores of almost every category was unaffected by respondents' experience level. However, we observed some important exceptions. Statistically significant dependencies ($p<0.05$) on expert experience level were found within the reproduction and adverse effects categories. In the reproduction category, we found statistically significant differences in mean scores for chronic median effective dose (ED_{50}) data ($p=0.0376$), and both acute and chronic NOEC data ($p=0.00508$ and $p=0.0023$, respectively). In the adverse effects category, both acute and chronic NOEC data ($p=0.0430$ and $p=0.0102$, respectively) were scored significantly based on experience level.

Discussion and Conclusions. The results of our SME survey provided results that parallel what has been thought to be a broad consensus in the ecotoxicological community, that observations related to chemical impacts on reproduction are of the greatest value for biological effects assessment. The goal of regulatory assessments for a variety of national and international regulatory bodies including the US EPA, Environment Canada, among others, is to maintain population sustainability (US EPA, 1997; Canadian Council of Ministers of the Environment, 1997; Bradbury et al., 2004). Measures of reproduction tend to be the most direct indicators of the potential to maintain viable populations, and thus it is intuitive that SMEs within the toxicological sciences, ecotoxicology, and risk assessment communities of practice find this class of information to be the most important regarding biological effects assessment.

It remains unclear whether an increase in the number of survey respondents would substantially alter the category score distributions. Figure 2 illustrates that all subclasses are distributed in a nearly uniform manner across the score interval. To better resolve the aggregate scores, it may be advantageous to address the number of type of intrinsic confounding factors contributing to each SME's opinion regarding a subclass of information. For example, and perhaps surprisingly, we found that mean scores for most categories were not significantly affected by self-reported experience level. While this result does not necessarily indicate that expert agreement is independent of experience level (note the large variance in score distributions, Figure 2), it suggests that mean scores may be only weakly affected by expert experience level. Although there was no significant effect from experience level on the majority of categories tested, it should be noted that the category scores significantly affected by expert experience level (acute and chronic NOEC data, see Results) were also among the categories that received the highest importance scores (Figure 2). Interestingly, among the reproduction category data, chronic ED_{50} data scored lowest among senior-level experts (i.e., early and mid-career experts scored these values, on average, much higher). Both acute and chronic NOEC data in the reproduction category scored higher among mid-career experts, followed by senior-level experts, while early-career experts scoring these data, on average, the lowest. Among the two significantly scored adverse effects information (i.e., acute and chronic NOEC data), mid-career and senior-level experts assigned similarly high scores, with early-career experts assigning consistently lower scores to these data. Unfortunately, we could not determine the cause of the disparities from experience level based solely from the reported scores.

While the curious observation of consistently large variance in SME opinion is intriguing, it is potentially not surprising. Even within the relatively narrow community of practice represented within this SME opinion survey, competing camps exist which hold differing viewpoints on the "appropriate" observations needed for an ecotoxicological assessment. For example, the possibility exists that researchers and regulators harbor biases regarding methods of inquiry and/or assessment practices that either consciously or subconsciously affect their scoring decisions. Regardless of competing influences, an agreement in opinion was observed across the overall community of practice, wherein chronic exposures and, generally, observations relating to reproduction, are considered high-value data for ecotoxicological assessment. This result suggests that individual bias may not entirely obscure the predominant trends of the field, and may therefore serve as a mechanism sufficient to identify consensus across a broadly trained expert population. However, a broader survey of SME opinion beyond DOD scientist and risk assessors would be of benefit to determine if the result of our scoping evaluation is maintained across the entire community of practice.

Scores among the SMEs did not appear to be completely debated. For example, some agreement in the scores appears likely for certain categories, such as ADME and non-adverse effects. Additionally, several subclasses can be seen to support a more localized spread than others in the same category, such as the acute benchmark dose subclass in the adverse effects category. In these cases, additional survey results could potentially lead to a convergence of consensus (by a reduction in the distribution variance), and therefore be worthy of additional investment. Strategies and best practices for increasing survey participation and reaching a wider audience include providing an email with a direct URL link to the survey with repeated email reminders, clearly stating that the survey results will be used (and how), assure that responses will be kept anonymous, provide opportunities for constructive criticism of the survey itself, and possibly providing rewards such as small prizes to incentivize participation (Nulty, 2008).

Subject matter experts were asked to evaluate both the lines of toxicological information as well as the major classes of ecotoxicological information in a generic context (i.e. score given no case-specific scenario). In certain case-specific scenarios, for example where thoroughly developed and validated *in silico* models may have excellent ability to predict toxicological outcomes of regulatory concern, scoring for *in silico* methods would score more highly. Therefore, we anticipate that the SME opinions gathered from our survey would look very different given case-specific scenarios, and we stress that the SME opinions that we have compiled in this study represent baseline opinions for use in a generalized context. Given the context-independent nature of the SME opinion data, the responses provide a useful snapshot of the perceived value connected to the various lines and classes of ecotoxicological data.

Applications and Future Work. Because our efforts constitute a scoping study, we recommend that future efforts focus on extending survey distribution to include a much broader cross-section of the overall community of practice. Expanded survey results could be used to specifically assess the similarities and differences in SME opinions of our representative survey of DOD scientists and risk assessors with opinions of academic, industrial, non-profit and non-DOD government organizations. Although we anticipate that the core valuation of ecotoxicological information for the purposes of risk assessment will be largely conserved across employment sectors, we highly recommend that this hypothesis be tested empirically.

Given the results of our SME opinion analysis, recommendations can be provided for prioritizing study types in support of ecotoxicological effects assessment on DOD facilities and can provide useful inferences to the overall ecotoxicology and environmental risk assessment communities of practice. Specifically, if resources were unlimited, the primary characterization of chemical effects would be facilitated through chronic reproductive as well as ADME effects assessments. In reality, these are expensive and difficult assessments to execute. These observations demonstrate a need for assay and method development that can provide predictive metrics of reproduction and ADME effects using more time and cost effective methods.

While non-adverse effects on individuals were considered to be the least important observations regarding the overall classes of biological information, toxicological and ecotoxicological studies tend to report all statistically significant effects found in a chemical exposure assessment as part of due diligence in scientific reporting. Significant non-adverse effects may be incorporated into lowest observed effect concentration (LOEC) and related NOEC values and can eventually be used as regulatory values if set as a toxicological benchmark. However, given SME opinion, this practice should be scrutinized to make certain that regulatory standards are not derived based on subtle effects that might not be of importance for animal fitness or population sustainability.

Finally, the SME opinion distributions can be useful to inform novel LCIA ecotoxicity characterization factors which incorporate multiple lines of toxicological evidence, as alluded to by Owens (1998). For example, lethal and sub-lethal dose-response relationships can be integrated via stochastic simulations drawing upon the virtual expert distributions as weights for different lines of evidence that can then be aggregated into a single ecotoxicity impact factor. Given stochastic concentration/exposure inputs, well-defined dose-response relationships, and SME opinion distributions as relative weighting factors, Monte Carlo methods may be applied to aggregate dissimilar, yet relevant, lines of toxicological information to more fully characterize the biological-effects potential (and related uncertainty) associated with LCA applications.

Acknowledgements. This research was supported by the US Army's 6.2 and 6.3 applied research program in Life Cycle Assessment. Opinions, interpretations, conclusions, and recommendations are those of the author(s) and are not necessarily endorsed by the U.S. Army. Kind thanks to the subject matter experts from the Tri-Services Toxicology Consortium and the US Army ERDC, Environmental Laboratory for making this study possible.

References

Ankley GT, Bennett RS, Erickson RJ, Hoff DJ, Hornung MW, Johnson RD, Mount DR, Nichols JW, Russom CL, Schmieder PK, Serrano JA, Tietge JE, Villeneuve DL. 2010. Adverse Outcome Pathways: A Conceptual Framework to Support Ecotoxicology Research and Risk Assessment. *Environmental Toxicology and Chemistry,* 29(3): 730-741.

Belton V, Stewart TJ. 2002. *Multiple Criteria Decision Analysis: An Integrated Approach*. Kluwer Academic Publishers: Boston.

Bradbury SP, Feijtel TCJ, Leeuwen CJV. 2004. Peer Reviewed: Meeting the Scientific Needs of Ecological Risk Assessment in a Regulatory Context. *Environmental Science and Technology*, 38(23):463A-470A.

Canadian Council of Ministers of the Environment. 1997. A Framework for Ecological Risk Assessment - Technical Appendices, Report. The National Contaminated Sites Remediation Program. Winnipeg, Manitoba, Canada. En 108-4/10-1-1997E. ISBN 0-660-16953-3.

Collier ZA, Bates ME, Wood MD, Linkov I. 2014. Stakeholder Engagement in Dredged Material Management Decisions. *Science of the Total Environment,* 496: 248–256.

Cooke RM, Goossens LHJ. 2008. TU Delft Expert Judgment Data Base. *Reliability Engineering & System Safety,* 93(5): 657-674.

Freedman D, Diaconis P. 1981. On the Histogram as a Density Estimator: L_2 Theory. *Probability Theory and Related Fields*, 57(4): 453-476.

Keeney RL. 2009. *Value-Focused Thinking: A Path to Creative Decision Making*. Harvard University Press: Cambridge.

Nulty DD. 2008. The Adequacy of Response Rates to Online and Paper Surveys: What can be Done? *Assessment & Evaluation in Higher Education,* 33(3): 301-314.

Owens JW. 1998. Life Cycle Impact Assessment: The Use of Subjective Judgments in Classification and Characterization. *International Journal of Life Cycle Assessment,* 3(1): 43-46.

Prado-Lopez V, Stewart T, Makowski M, von Winterfeldt D. 2013. Value Measurement Analysis of Energy Tradeoffs in South Africa. *Proc. ISSST, v1*, http://dx.doi.org/10.6084/m9.figshare.805096.

Prado-Lopez V, Seager TP, Chester M, Laurin L, Bernardo M, Tylock S. 2014. Stochastic Multi-Attribute Analysis (SMAA) as an Interpretation Method for Comparative Life-Cycle Assessment (LCA). *International Journal of Life Cycle Assessment,* 19: 405-416.

US EPA .1997. Ecological Risk Assessment Guidance for Superfund Process for Designing and Conducting Ecological Risk Assessments - Interim Final. EPA 540-R-97-006, OSWER 9285.7-25, PB97-963211, pp. 1-28.

Supplementary Information

Ranking the Relative Importance of Toxicological Observations Based on Subject Matter Expertise

Kurt A. Gust US Army Engineer Research & Dev. Center, Kurt.A.Gust@usace.army.mil
Michael L. Mayo US Army Engineer Research & Dev. Center, Michael.L.Mayo@usace.army.mil
Zachary A. Collier US Army Engineer Research & Dev. Center, Zachary.A.Collier@usace.army.mil

Appendix A: Supplementary Information on SME Survey.

The 6 Major Classes of Biological Information. Each of the six major classes of biological information were incorporated into the subject matter expert survey to allow the interviewees to score the importance of each within ecotoxicology and risk assessment.

1. *Lethality* – Lethality represents a quintessential effect in toxicology. A variety of assays and statistical models have been developed to characterize the lethal toxicity of chemical exposure. The subject matter expert survey incorporated each of the major summary categories for lethality bioassays including: (A) **LC**50 and **LD**50 which describe the **C**oncentration **or D**ose causing 50% lethality in a sample population at a given exposure duration, (B) **N**OEC and **L**OEC which describes the **N**o observed effects concentration or **L**owest observed effects concentration for a chemical, (C) BMDL which represents the bench mark dose level which is associated with the statistical point of departure from the control state given a dose response curve, and finally (D) the species sensitivity distribution provides a summary of lethality information for multiple species allowing sensitivity comparisons and a relative range of toxic doses. An additional component that is critical to the interpretation of lethality information is the length of chemical exposure. Short term exposures are classified as “acute” exposures in toxicological research whereas extended exposures that persist for a long duration of the species life cycle are termed “chonic” exposures. Within the survey, we allowed interviewees to score both acute and chronic exposure durations, not only for lethality, but also for all other relevant classes of biological information (Figure S1).

2. *Sublethal Effects* – All non-lethal effects on individuals were summarized as sublethal effects. Sublethal effects were broken into three major categories: (1) Effects on Reproduction, (2) Adverse Effects on Individuals and (3) Non-Adverse Effects on Individuals. Similar to assessment of lethality a variety of assays and statistical models have been developed to characterize the effects of chemical exposure. The subject matter expert survey incorporated each of the major summary categories for sublethal effects results: (A) **EC**50 and **ED**50 which describe the **C**oncentration or **D**ose causing 50% of a prescribed **E**ffect in a sample population at a given exposure duration, (B) NOEC and LOEC, (C) BMDL and (D) the species sensitivity distribution for a given sublethal effect (Figure S1).

3. *Absorption, Distribution, Metabolism & Elimination (ADME)* – In order for a chemical to elicit a toxicological effect, it must first be accumulated in the body. Information about how the contaminant is absorbed into the body, the distribution of the contaminant throughout the body, metabolic action on the contaminant (i.e. contaminant transformation) and elimination of the contaminant from the body (ADME) has great utility in toxicology for explaining chemical toxicity.

ADME is represented as five parameters in the SME survey: (1) Critical tissue residue which represents the tissue concentration that causes a prescribed effect, (2) Octanol water partitioning coefficient (KoW) which is a non-biological method that is used to estimate the bioaccumulation potential of a chemical based on its physicochemical properties, (3) bioconentration factor (BCF) which represents an empirical characterization of the potential for a chemical to concentrate in organisms, (4) toxico-kinetic and toxico-dynamic parameters representing a composite of empirical observations of chemical uptake / elimination kinetics as well as metabolic rates and products of chemical transformation, and (5) biomagnification potential representing the likelihood that chemical concentrations in tissue will be magnified as it moves up the food chain (Figure S1).

4. *Systems Biology* – The systems biology category represents the integration of molecular toxicology to toxicological impacts in the organism and or populations. Systems biology is represented by four parameters including: (1) mode of action (MOA), which represents the specific type of toxicity elicited by a chemical or chemical class (i.e. neurotoxicity versus hepatotoxicity), (2) molecular pathway impacts which is a subset of (MOA) represented by specific metabolic pathways that are perturbed by a chemical, (3) mutation and carcinogenesis, which characterizes the potential for a chemical to cause DNA mutations and/or cause cancer, and (4) adverse outcome pathways, which are integrate toxicological information across levels of biological organization from molecular initiating events to adverse outcomes in the individual / population (Figure S1).

5. *Population-level effects* – Demonstration that a chemical can affect populations or population sustainability are summarized as population-level effects. Population-level effects are summarized as three categories: (1) effects on intrinsic population growth representing the maximum population growth rate free of density-dependent forces, (2) reduced genetic diversity of the population and (3) population demography comparing exposed field sites to reference (un-contaminated) field sites (Figure S1).

6. *Community-level effects* – Demonstration that a chemical can affect the composition of a biological community are summarized as community-level effects. Community-level effects are represented by two classical metrics of community structure: (1) species diversity represents the number of species found in an environment coupled with the evenness of individuals represented within each species and (2) species richness representing the simple count of species in a given environment (Figure S1).

The 3 Lines of Biological Information. Within the six major classes of biological information, there are three potential sources or “lines” from which data may be derived: *in vivo*, *in vitro* and *in silico* observations. In toxicological research *in vivo* observations are derived from experiments which utilize whole animal exposures where toxicological effects are observed in individuals or groups of individuals. *In vitro* toxicological research leverages cultured tissues or groups of cells surviving outside of a whole organism system providing toxicity observations for tissues, cells / cellular functions and/or molecular effects. Finally, *in silico* observations leverage computational models to generate predictions of toxicological effects based on prior knowledge from related chemical types (i.e. quantitative structure-activity relationships) or other known cause and effect relationships between chemical exposure and toxicological effects.

Participant Information. To gauge the expertise and level of experience for each SME, a request for demographic information was additionally included in the survey. Requested demographic information included: title, position / rank, subject disciplines (primary), subject disciplines (secondary) and years of experience.

Survey for Determining Importance of Ecotoxicological Parameters for ECO-Life Cycle Assessment

Rank Lines of Information

(Rank 1-3 with 1 being first choice)

Rank	Line of information
	in silico data
	in vitro data
	in vivo data

Interviewee Info:

Field	Entry
Title (*Not Your Name Please*)	
Position / Rank	
Subject Discipline(s) (Primary)	
Subject Discipline(s) (Secondary)	
Years of Experience	

Rate Importance of Toxicological Categories (Pure Benefit Metric - Score All Categories Independent of Cost)

Rating Scale 0 - 100 (0 = Not Important, 100 = Critically Important). No math needed, every box can be 0-100.

Score [ ] **Lethality** *(chemical exposure causes death)*

Acute Score	Chronic Score	Parameter
		LD50 / LC50 (**D**ose or **C**oncentration causin 50% **L**ethality in a sample population)
		NOEC or LOEC (No Observed Effect Concentration / Lowest Observed Effect Concentration)
		BMD (Bench Mark Dose)
		Species Sensitivity Distribution

Score [ ] **Sublethal Effects**

[ ] **Reproduction** *(chemical exposure causes impairments in reproduction)*

Acute Score	Chronic Score	Parameter
		ED50 / EC50 (**D**ose or **C**oncentration causin 50% of an **E**ffect in a sample population)
		NOAEC / LOAEC (No Observed Adverse Effect Concentration / Lowest Observed Adverse Effect Concentration)
		BMD (Bench Mark Dose)
		Species Sensitivity Distribution

Score [ ] **Adverse Effects on Individual** *(all adverse impacts **excluding** impacts on reproduction)*

Acute Score	Chronic Score	Parameter
		ED50 / EC50 (**D**ose or **C**oncentration causin 50% of an **E**ffect in a sample population)
		NOAEC / LOAEC (No Observed Adverse Effect Concentration / Lowest Observed Adverse Effect Concentration)
		BMD (Bench Mark Dose)
		Species Sensitivity Distribution

Score [ ] **Non-Adverse Effects on Individual** *(statisitcally significant change, however no obvious fitness implications, ie. subtle behavioral changes)*

Acute Score	Chronic Score	Parameter
		ED50 / EC50 (**D**ose or **C**oncentration causin 50% of an **E**ffect in a sample population)
		NOEC / LOEC (No Observed Effect Concentration / Lowest Observed Effect Concentration)
		BMD (Bench Mark Dose)
		Species Sensitivity Distribution

Score [ ] **Absorption, Distribution, Metabolism & Elimination (ADME)**

Acute Score	Chronic Score	Parameter
		Critical Tissue Residue

Score	Parameter
	Kow (Octanol Water Partitioning Coefficent)
	Bioconcentration Factor (BCF)
	Toxico-Kinetic & Toxico-Dynamic Parameters
	Biomagnification Potential

Score [ ] **Systems Biology** *(molecular, genetic and metabolic impacts linked to chemical toxicity)*

Acute Score	Chronic Score	Parameter
		Mode of Action
		Mechanism of Action
		Molecular Pathway Impacts
		Mutation & Carcinogenesis
		Adverse Outcome Pathways

Score [ ] **Population-Level Effects** *(effects of chemical on population-level parameters)*

Score	Parameter
	Intrinsic Population Growth Rate (Lambda)
	Reduced Genetic Diversity in Population
	Population Demography - Exposed Field Site vs. Reference Field Site

Score [ ] **Community-Level Effects** *(effects of chemical on population-level parameters)*

Score	Parameter
	Species Diversity - Exposed Field Site vs. Reference Field Site
	Species Richness - Exposed Field Site vs. Reference Field Site

Figure S1: Subject Matter Expert Survey. Questionnaire distributed to subject matter experts.