Data Evaluation Criteria for Environmental Hazard Studies Under TSCA

Amelia Nguyen[1], Sarah Au[1], Karen Eisenreich[1], Colleen Elonen[2], Dale Hoff[2], Tracy Wright[1], and Amuel Kennedy[1]
[1]U.S. EPA, OCSPP/OPPT, Washington, DC
[2]U.S. EPA, ORD/CCTE/GLTED, Duluth, MN

Amelia Nguyen | Nguyen.Amelia@epa.gov | 202-564-4268

Systematic Review Process

Systematic Review Process: The U.S. EPA's Office of Pollution Prevention and Toxics (EPA/OPPT) intends to apply systematic review (SR) in developing risk evaluations under TSCA**. Systematic Review is a comprehensive, unbiased, transparent and reproducible way to identify relevant literature on a topic.** TSCA Section 6(b)(4)(F) states that EPA shall "integrate and assess available information on hazards and exposures" and "describe the weight of scientific evidence." Sections 26(h) and (i) require that EPA employ best available science and make decisions based on the weight of the scientific evidence. The Risk Evaluation rule defines weight of the scientific evidence as a "systematic review method, applied in a manner suited to the nature of the evidence or decision, that uses a pre-established protocol to comprehensively, objectively, transparently, and consistently, identify and evaluate each stream of evidence, including strengths, limitations, and relevance of each study and to integrate evidence as necessary and appropriate based upon strengths, limitations, and relevance."

Introduction: OPPT is conducting SR to evaluate environmental toxicological data for high priority existing chemicals undergoing risk evaluation under the TSCA. The TSCA SR includes several steps, from protocol development to evidence integration (**Figure 1**). It is integral for the TSCA Risk Evaluations (RE), which are fit-for-purpose, to be based on the best available science and a weight of the scientific evidence process (U.S. EPA, 2018). To identify acceptable on-topic references for title and abstract, and full-text screening of environmental hazard data, EPA relied on the process used to populate the ECOTOX database (U.S. EPA, 2016a and www.epa.gov/ecotox) and backwards searching approaches for key supporting gray literature (**Figure 2**). Following full-text screening, environmental hazard data were extracted and evaluated using predefined criteria. EPA developed data quality evaluation criteria under TSCA, based on a combination of EPA's ECOTOX criteria (U.S. EPA, 2016a) and the Criteria for Reporting and Evaluating ecotoxicity Data (CRED) (Moermond et al., 2016), to determine overall study quality prior to the start of the evaluation step. EPA implemented quality controls to promote a consistent and transparent evaluation process.

Objectives: This poster describes the Data Evaluation process for assessing the quality of multiple data types supporting the environmental hazard assessment. The data resulting from OPPT's SR analysis will serve as the basis for hazard characterization and will be integrated into the risk evaluation for each of the identified high priority substances under TSCA (U.S. EPA, 2016b and EPA, 1998). The environmental hazard evaluation criteria are being updated to better consider methodological design, implementation, and reporting to address public and peer review comments.

ECOTOX Pipeline within Systematic Review Under TSCA

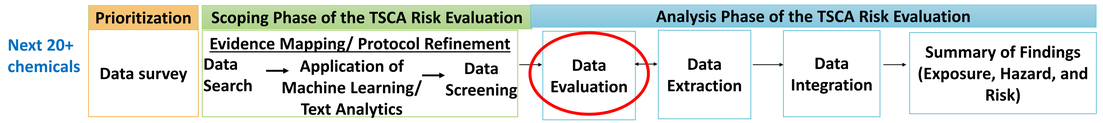

Figure 1. Flow diagram of TSCA SR pipeline.

Key Terms in Data Evaluation

- **Domain**: The general categories of data/information attributes intended to assess methodological conduct and risk of bias
- **Metric**: The sub-categories of domain attributes
- **Criteria**: Specific criteria are developed for each metric, which express conditions of the quality level assigned to the metric (high, medium, low, or unacceptable)
- **Data Quality Score :** Quantitative score calculated following evaluation of discipline-specific and data type-specific data evaluation domains and metrics according to predefined scoring criteria and accounting for metric weighting factors

ECOTOX Literature Search and Study Selection Flow Diagram

A comprehensive search and review of toxicity data in open and gray literature (e.g., government documents) was conducted, with transparent standard operating procedures that meet requirements for systematic review.

- Data were extracted for >200 fields of information (see **Table 1**).
- Streamlines the cost for literature searches and data curation within the Agency and provides all information in public format for States, Tribes, Industry, and International governmental entities.
- Continuous update of protocols and annual evaluation of most applicable sources to ensure the inclusion of relevant publications.

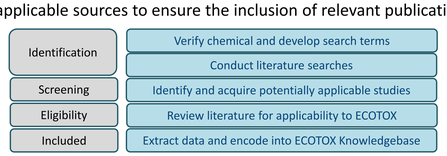

Figure 2. Flow diagram of ECOTOX pipeline.

Data Types/Extraction Fields in ECOTOX

Category	Data Fields (not all inclusive)	Select study evaluation questions with relevant ECOTOX field(s)
Chemical	Chemical Name, CASRN, Grade, Purity, Formulation, Carrier *Test Specific:* Analysis, Application Type and Rate/Frequency, Number of Doses, Doses, Concentration Type (e.g., active ingredient or formulation), Concentration/Dose associated with each effect and/or endpoint	Is test substance identified? **Required for inclusion in ECOTOX** Is the purity of test substance reported? Chemical Purity Were chemical concentrations verified? Chemical Analysis (e.g., nominal versus measured concentrations)
Species	Scientific and Common Name, Taxonomy, Lifestage, Age, Initial and Final Weight, Gender, Source	Is the species given? **Verifiable species (Scientific Name, etc.) required for inclusion in ECOTOX** Are the organisms well described? Organism Source, Lifestage, Age, Gender, Initial and Final Weight
Test Conditions	Test Method, Media Type, Test Location, Exposure and Study Duration, Control, Experimental Design, Physical and Chemical Soil and Water Parameters	Are appropriate controls performed? **A control is required for inclusion in ECOTOX**, type described in Control Is a guideline method (e.g., OECD) used? Test Method Are the experimental conditions appropriate and acceptable for the test substance and organism? Test Method, Media Type, Test Location, Experimental Design, Physical and Chemical Soil and Water Parameters (e.g., pH, Temperature, Dissolved Oxygen)
Test Results	Effect (observation of a response): general effect groups and specific effect measurements, Endpoint (quantification of an observed effect, e.g., LC50), Trend, Response Site, Effect %, Statistical Significance and Level, Observed Duration (exposure Duration when result observed), Bioconcentration (BCF or BAF) with units	Are the reported effects and endpoints appropriate for the purpose, test substance and organism? Effect Measurement, Endpoint Is the response/effect statistically significant? Statistical Significance, Significance Level

Table 1. Types of data, as reported by authors, extracted from each reference with category, data fields, and examples of how ECOTOX fields can inform study evaluation questions.

TSCA Environmental Hazard Data Evaluation Process

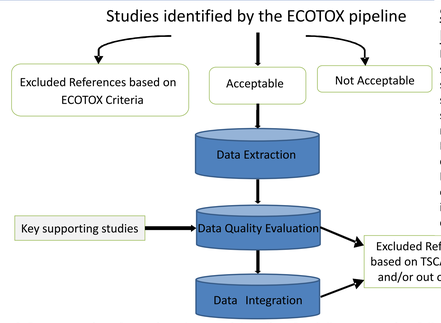

Figure 3. Data extraction of aquatic and terrestrial studies after full-text screening, followed by data quality evaluation and integration. Data extraction tables may change depending on inclusion/exclusion of pathways due to regulatory nexus.

Step 1. DATA EXTRACTION

Data gathered for the high priority substances under TSCA were subjected to title and abstract screening and then full-text screening using the ECOTOX methodology and criteria (U.S. EPA, 2016a and 2018). Acceptable data endpoint values (e.g., LC/EC_{50}s, NOAELs, etc.) were extracted into spreadsheets for use in problem formulation and the data integration phase (**Figure 3**).

Evaluation Domain	Metric
Test Substance	Metric 1: Test Substance Identity Metric 2: Test Substance Source Metric 3: Test Substance Purity
Test Design	Metric 4: Negative Controls Metric 5: Negative Control Response Metric 6: Randomized Allocation
Exposure Characterization	Metric 7: Experimental System/Test Media Preparation Metric 8: Consistency of Exposure Administration Metric 9: Measurement of Test Substance Concentration Metric 10: Exposure Duration and Frequency Metric 11: Number of Exposure Groups and Spacing of Exposure Levels Metric 12: Testing at or Below Solubility Limit
Test Organisms	Metric 13: Test Organism Characteristics Metric 14: Acclimatization and Pretreatment Conditions Metric 15: Number of Organisms and Replicates per Group Metric 16: Adequacy of Test Conditions
Outcome Assessment	Metric 17: Outcome Assessment Methodology Metric 18: Consistency of Outcome Assessment
Confounding/Variable Control	Metric 19: Confounding Variables in Test design and Procedures Metric 20: Outcomes Unrelated to Exposure
Data Presentation and Analysis	Metric 21: Statistical Methods Metric 22: Reporting of Data Metric 23: Explanation of Unexpected Outcomes

Table 2. Environmental Hazard Data Evaluation Domains and Metrics (U.S. EPA, 2018).

Step 2. DATA QUALITY EVALUATION

EPA developed 23 data quality evaluation criteria for environmental data. The criteria fall under seven Domains: Test Substance, Test Design, Exposure Characterization, Test Organisms, Outcome Assessment, Confounding/Variable Control and Data Presentation and Analysis (**Table 2**) (U.S. EPA, 2018). For each metric, quality levels are based on strengths and limitations. Metrics are scored as high, medium, low or unacceptable (1, 2, 3 or 4). Depending on the study, some metrics may not be scored and rated as N/A. Metrics are assigned weighting factors: critical metrics = 2, other metrics = 1. Critical metrics were chosen based on 1) professional judgment and 2) greatest ability to inform hazard identification. Metric scores are multiplied by weights and combined for an overall study quality score. Scores are not intended to imply precision and/or accuracy of scoring results. Reviewers can adjust the overall quality to capture professional judgment, if appropriately justified. This numerical scoring method is used to convert the quality level for each metric into the overall quality level for the data/information source. A study is disqualified from further consideration if the quality level of one or more metrics is rated as unacceptable (i.e., score of 4).

Scoring System: Metrics for Study Quality

High: No notable deficiencies or concerns are identified in the domain metric that are likely to influence results.
Medium: Minor uncertainties or limitations are noted in the domain metric that are unlikely to have a substantial impact on results.
Low: Deficiencies or concerns are noted in the domain metric that are likely to have a substantial impact on results.
Unacceptable: Serious flaws are noted in the domain metric that consequently make the data/ information source unusable.
Not rated/ applicable: Rating of this metric is not applicable to the data/information source being evaluated. Not rated/applicable will also be used in cases in which studies cite a literature source for their test methodology instead of providing detailed descriptions. In these circumstances, EPA will score the metric as Not rated/Applicable and capture it in the reviewer's notes. If the data/information source is not classified as "unacceptable" in the initial review, the cited literature source will be reviewed during a subsequent evaluation step and the metric will be rated at that time.

Updates to the Environmental Hazard Data Quality Criteria

The 23 criteria for evaluating environmental hazard data have been updated to increase transparency and consistency regarding metric descriptions and weight for different metrics used to calculate the data quality score. Specifically, OPPT has:

- Established a process where at least two assessors will evaluate the same study and arbitration meetings/emails were carried out to resolve conflicting evaluations to improve evaluation consistency across chemicals and assessors.
- Evaluated chemical-specific case nuances that may have resulted in inconsistent metric evaluations.
- Clarified the descriptions of the metrics, especially for the N/A ratings. For example, Metric 12 (testing at or below the solubility limit): physical particles (e.g., asbestos fibers) or and specific exposure pathways used in the study (e.g., diet, sediment, soil)-should be deemed N/A.
- Revised the list of critical metrics that have a heavier weighting factor. EPA is proposing to downgrade the weight of Metrics 13 (Test Organism Characteristics), 17 (Outcome Assessment Methodology), and 22 (Reporting of Data) to a weight of 1 in the environmental hazard data quality evaluation criteria.
- Implemented best practices to ensure reviewers' comments are focused on the strengths and/or limitations of the study element(s) considered in the metrics for each paper.

Future Directions

- Further refine individual metrics as they relate to quality levels (e.g., what rating should be applied when the information is lacking).
- Evaluate whether current unacceptable ratings are appropriate.
- Consider whether a single unacceptable metric rating should make the full study unacceptable.
- Consider whether EPA should continue using the numerical scoring approach (**Figure 4**).

Overall Score=
$$\sum(\text{Metric Score} \times \text{Weighting Factor}) \div \sum(\text{Metric Weighting Factors})$$

High	Medium	Low	Range of Overall Score:
≥1 and < 1.7	≥1.7 and <2.3	≥2.3 and ≤3	**1 to 3**

Figure 4. Overall score calculations.

- Collapse high, medium, low, not rated, or unacceptable quality bins for some metrics.
- Incorporate quality control measures involving two reviewers and conflict arbitration.
- Seek feedback regarding the updates made to the Environmental Hazard Data Quality Criteria, metric descriptions, and how metrics could be evaluated.
- Publish a Generic Systematic Review Protocol.

REFERENCES

1. Moermond et al. 2016. CRED: Criteria For Reporting And Evaluating Ecotoxicity Data. *Env Tox & Chem* 35 (5): 1297-1309.
2. U.S. EPA (U.S. Environmental Protection Agency). 2018. *Application of Systematic Review in TSCA Risk Evaluations*. (EPA/740/P-1/8001). Washington, DC: U.S. Environmental Protection Agency, Office of Chemical Safety and Pollution Prevention.
3. U.S. EPA. 1998. Guidelines for Ecological Risk Assessment. (EPA/630/R-95/002F). Washington, DC: U.S. Environmental Protection Agency, Risk Assessment Forum.
4. US EPA. 2016a. ECOTOX Knowledgebase. Cited April 6, 2020. https://cfpub.epa.gov/ecotox/help.cfm
5. U.S. EPA. 2016b. Weight of Evidence in Ecological Assessment. (EPA/100/R-16/001). Washington, DC: U.S. Environmental Protection Agency, Risk Assessment Forum.

Acknowledgement: We thank Sidrah Khan for her collegial discussions and advice and help in creating this poster.