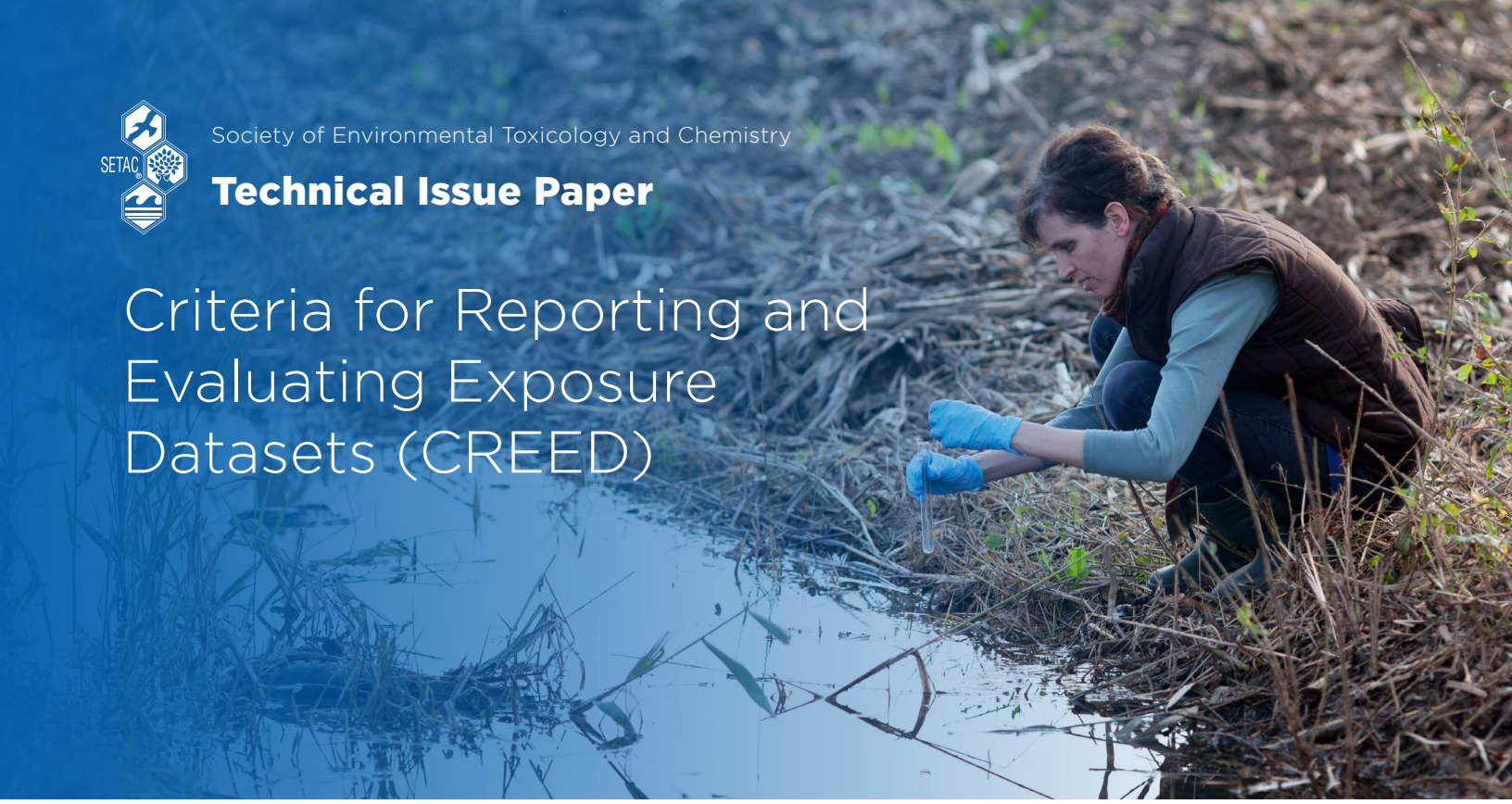

Society of Environmental Toxicology and Chemistry

Technical Issue Paper

Criteria for Reporting and Evaluating Exposure Datasets (CREED)

Background

Environmental risk assessment of chemicals determines the nature and likelihood of harmful effects occurring to organisms. Ecological risk assessments generally rely on two key values: chemical exposure and chemical hazard potential, which are typically generated from chemical monitoring data and ecotoxicity data, respectively.

Comprehensive programs exist to determine the suitability of ecotoxicity data for specific assessment purposes but analogous comprehensive schemes have thus far not been available for evaluating exposure data.

The usefulness of a risk assessment hinges on the suitability of the exposure data used in terms of its relevance and reliability.

- Reliability refers to the inherent quality of the dataset, and its evaluation focuses on the methods used for sample collection, chemical analysis, and data processing and statistics.
- Relevance refers to the degree of suitability of an existing environmental monitoring dataset, or one that will be generated, to address the specific purpose as defined by the assessor.

In the absence of a systematic and transparent approach to evaluating exposure data, differences may occur in how individual assessors evaluate and use environmental chemistry data, potentially leading to discrepancies or disagreements about chemical prioritization and wide-scale environmental risk.

What Is CREED?

Criteria for Reporting and Evaluating Exposure Datasets (CREED) is a practical tool that provides a systematic approach and criteria for the consistent and transparent evaluation of the reliability (quality) and relevance (fitness for purpose) of exposure data for use in risk assessment. The CREED approach applies to many different purposes of assessment that utilize exposure data. CREED can be used by risk assessors to evaluate existing datasets or to identify knowledge gaps for use in planning future monitoring studies. CREED also can be used by data generators as a guide to the parameters that should be collected and reported and by database owners as a guide for which data fields are considered important to include in the database—in both cases so that their data can be used, even by outside researchers, to support environmental assessments.

CREED was initiated as a SETAC-supported activity with the goal of improving the transparency and consistency with which exposure data are evaluated for use in environmental assessments. As such, CREED provides a framework through which expert judgement is guided and appropriately documented so that the evaluation is

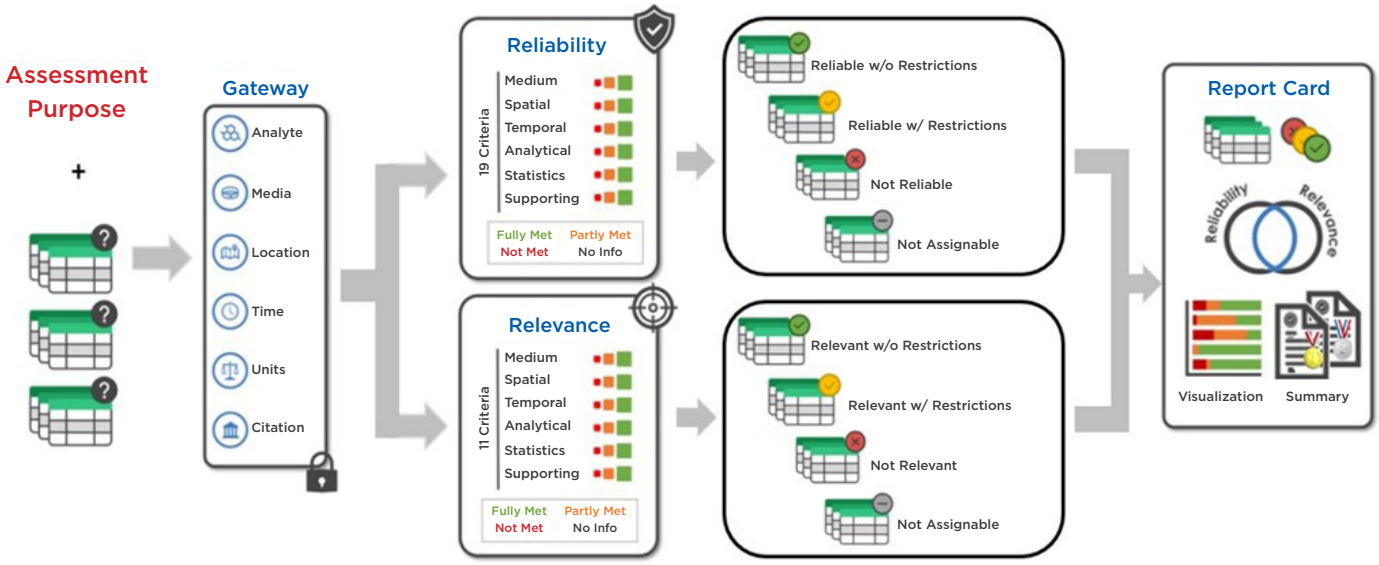

Figure 1. Outline of CREED Approach

transparent, systematic, and consistent from user-to-user. Figure 1 provides an outline of the CREED approach.

CREED Steps

Assessment Purpose Statement

When utilizing CREED, scientists are first directed to prepare a purpose statement. A specific, detailed purpose statement will increase the objectivity and transparency of the relevance evaluation—both of which are key elements of the CREED approach.

Whereas reliability is generally independent of the purpose for which data are being assessed, evaluating relevance is naturally dependent on the purpose of the study in which the data are being considered for use. Therefore, the data evaluation process starts with a dedicated purpose statement, which ideally will specify the types and amounts of information that is required for the particular risk assessment being performed.

Gateway Criteria

In the second step of CREED, a given dataset is first compared to six "gateway" criteria, which define the minimum amount of information required to evaluate a study. This is a time-saving step so an assessor can quickly eliminate datasets that are missing basic information: the chemical analyzed, sampling medium, sampling location, sampling year, units of measurements, and a citation. If a dataset fails one or more gateway criteria because of missing information, then the dataset is dropped from further consideration, unless the missing data can be located. Datasets that pass the gateway stage then undergo detailed evaluation for reliability and relevance.

Detailed Reliability and Relevance Criteria

The evaluation for datasets that pass the gateway stage is more rigorous. There are 19 reliability criteria and 11 relevance criteria that must be assessed before scoring.

The reliability criteria address whether detailed information is reported about the analyte identity, sampling medium, sample collection method and handling, dates and times of sampling, and analytical detection limits. Some criteria address method performance and quality control, or whether data processing and statistics were handled appropriately. A common issue with datasets for trace organic chemicals is the presence of nondetects (also called censored data), which if inappropriate statistical techniques are used, can result in biased or flawed results; therefore, two criteria address censored data issues and statistics. The objective, if all reliability criteria are met, is that the assessor knows enough about the study design to understand what the data represent and can evaluate whether the method is appropriate for the analyte and sampling medium. For example, for a hydrophobic chemical, such as polychlorinated biphenyls (PCB), the criterion that addresses sample handling is important because PCBs in water will largely be associated with particulates; therefore,

filtration of water samples would result in loss of PCBs and would be inappropriate for this chemical for most purposes.

The relevance criteria address similar topics (analyte identity, sample medium, number of sampling sites, sampling frequency and duration, analytical sensitivity), but here, the focus is on whether these aspects of the study are appropriate for the specified assessment purpose. For example, a study of a chemical in estuarine water samples might be a reliable study, but the sampling medium would be inappropriate if the assessment purpose were to assess potential human health effects from drinking water consumption.

For each reliability and relevance criterion, a given study is rated as either:

- **Fully met**: All conditions described by the criterion are satisfied by the study;
- **Partly met**: Some of the conditions described by the criterion are met for either part or all of the dataset, or all conditions are met for part of the dataset;
- **Not met or inappropriate**: The data or approach is flawed or inappropriate for the purpose;
- **Not reported**: Insufficient information was provided to evaluate the criterion; or
- **Not applicable**: For circumstance-specific criteria only, the circumstances described by the criterion do not apply to the dataset.

Whenever a criterion is rated as "partly met," "not met or inappropriate," or "not reported," it is very important that the assessor record the data limitations, flaws, or missing information that triggered this rating. This is a key element of CREED because this provides information on data gaps that may restrict data use or be useful for estimating uncertainty in the risk assessment.

Scoring System

Ultimately, the study or dataset is assigned to each category in both reliability and relevance, based on the ratings of their respective criteria:

Relevance and Reliability Categories

- Reliable or relevant without restrictions,
- Reliable or relevant with restrictions,
- Not reliable or relevant, or
- Not assignable.

The dataset reliability and relevance categories are then combined to determine the overall usability of the dataset for the given purpose as usable without restrictions, useable with restrictions, or not usable.

The overall rating of the reliability and relevance criteria is done using a two-level scoring system—at silver and gold levels. This two-level scoring system was developed recognizing that "perfect" (gold standard) datasets are not common, while potentially usable datasets (i.e., those that meet basic criteria) are frequently encountered and should therefore be within the scope of CREED applicability for the framework to be of practical use. If only perfect datasets were allowable, then CREED would not be very useful. Therefore, CREED distinguishes between "required" criteria (which are important for most assessment purposes) and "recommended" criteria (which are considered less critical). At the silver score level, the dataset is scored based only on the required criteria, whereas all criteria (required plus recommended) determine the dataset at the gold level. Thus, the silver level is less ambitious than the gold level, which represents an ideal dataset.

- Gold: Includes recommended and required criteria
- Silver: Includes required criteria only

CREED Dataset Scoring

Every dataset is scored at both gold and silver levels, which are more and less rigorous, respectively. As an example, suppose that Dataset A fully met all required criteria, but had one or more recommended criteria rated as "not reported" because of missing information. Dataset A would be scored as "useable without restrictions" at the silver level, but "not useable" because of missing information at the gold level. If the missing information could be located, Dataset A might be re-evaluated and receive a higher useability score at the gold level.

The CREED Tool: Workbook and Report Card

To allow users to readily follow the workflow, evaluate, and create a CREED summary (report card) for any given dataset, CREED has been implemented in a Microsoft Excel® workbook template. Each step of the CREED approach is represented in this workbook. After the assessor enters ratings for the reliability and relevance criteria, the scoring tool automatically assigns the dataset to categories of reliability, relevance, and usability at both the silver and gold levels. In addition to these categories, an important product of the CREED

evaluation is a description of any identified data limitations (including missing information) that might restrict use of the dataset. Such data limitations also can guide the user by identifying missing information that, if located, would allow the user to more fully evaluate the dataset for use in the specified assessment purpose. The full evaluation outcomes can be exported as a downloadable report card.

Best Practices

CREED not only supports informed decision-making but also can be valuable in identifying data gaps and planning for future data collection campaigns.

For the chemical management assessment practitioner, CREED provides an approach and tools to consistently and transparently evaluate the quality of a chemical dataset and whether it is fit for a specific assessment purpose, as well as to identify limitations of the dataset that may qualify or constrain the use of the data.

For the data generator, CREED provides guidance on important study characteristics that should be reported to ensure that their data are useful to the widest possible range of assessment types.

For database managers, CREED can serve as guidance on the types of information about a dataset that are important for databases to include so that users can extract the information (metadata) they need about a dataset to fully evaluate that dataset for use in environmental assessments.

For the scientific community, CREED offers an opportunity to improve the quality and suitability of monitoring data that are used in environmental assessments.

Challenges

Challenges to the use of CREED include:

- Implementation of CREED will require the combined efforts of data generators, data users, and database owners to succeed.
- Over time, it is hoped that standard practice may be improved as critical data reporting steps are more widely adopted. In the meantime, the silver scoring system can serve as a practical compromise.
- CREED has been developed to address surface water, soil, sediment, and biota. Possible future expansion may include additional criteria applicable to nontargeted analysis, chemicals in air, and passive sampling.

Path Forward

In the interest of better science and decision-making, and to promote cost-effective environmental monitoring and assessment, the developers of CREED hope to make future improvements to CREED as the scientific community gains experience in the application of this approach. We are seeking feedback on its usability and any enhancements that can be added to the approach to improve its functionality. Reach the developers at science@setac.org.

Resources

www.setac.org/data-usability

Acknowledgements

SETAC is grateful for the efforts and contributions of everyone involved in the advancement of this topic, including all participants and sponsors (Concawe, GSK, Metals Environmental Research Association [MERA], SETAC Journals [ET&C and IEAM], Syngenta and Unilever) of the May 2022 SETAC technical workshop to develop CREED.

Copyright