

# Instructions for reliability and relevance evaluation of **ecotoxicity** studies using the SciRAP tool

For questions, please contact Marlene Ågerstrand, [marlene.agerstrand@aces.su.se](mailto:marlene.agerstrand@aces.su.se).

Reliability and relevance are equally important in the assessment of chemicals. It is therefore recommended that reliability as well as relevance are evaluated according to the process described below.

At [www.scirap.org](http://www.scirap.org), go to the tab “Ecotoxicity studies” at the top of the page, click on “Evaluate reliability & relevance ” or “Evaluate reliability & relevance - nano”. The reliability/relevance evaluation is conducted in two steps:

- 1) Evaluate the study according to the reliability and relevance criteria (including weighing and deselection of criteria), and
- 2) Assign the study to reliability and relevance categories (optional).

## **1. Evaluate the study according to the reliability and relevance criteria (including weighing and deselection of criteria)**

Fill out how well each criterion is met by choosing an alternative from the drop-down menu to the right of each criterion. Choose between “Fulfilled”, “Partially fulfilled”, “Not fulfilled”, “Not reported” and “Not determined”. Guidance from Moermond et al. (2016) and Hartmann et al. (2017) (for nanomaterials) is provided by clicking the question mark next to each criterion. Motivations and notes can be added in the comments fields.

Certain criteria may be considered more or less critical for the evaluation of specific types of studies or assessments. The criteria have all been given a default weight and this can be adjusted up by clicking “Increase weight” for criteria that are considered to be specifically important. Criteria that are not applicable for the specific study or question being assessed may be removed from the evaluation by clicking “Remove”. Motivations for weighing and removing criteria can be provided in the comments fields.

Note that your selections for the criteria will be saved on the computer until you click "Reset form".

Once the reliability and relevance evaluation is completed, name your file and click the green "Export to excel" button. An excel sheet will open up, illustrating the results of your evaluation. This excel sheet can be saved to your computer and shared with others. The excel file will provide summaries of the evaluation, including a colour profile in a bar chart showing green for “Fulfilled” criteria, yellow for “Partially fulfilled”, red for “Not fulfilled”, dark grey for “Not reported”, and light grey for “Not determined”. The bar chart takes the weights of criteria into account and does not include removed criteria. The colour profile gives an overview of the evaluation result, however, it does not dictate reliability/relevance category.

## 2. Assign the study to reliability and relevance categories (optional!)

After the study has been evaluated according to the criteria it may be assigned into one of the following reliability/relevance categories using expert judgment and the guidance provided in Moermond et al. (2016) and Hartmann et al. (2017) (for nanomaterials).

### a. Reliability categories

- *Reliable without restrictions:* All critical reliability criteria for this study are fulfilled. The study is well designed and performed, and it does not contain flaws that affect the reliability of the study.
- *Reliable with restrictions:* The study is generally well designed and performed, but some minor flaws in the documentation or setup may be present.
- *Not reliable:* Not all critical reliability criteria for this study are fulfilled. The study has clear flaws in study design and/or how it was performed.
- *Not assignable:* Information needed to make an assessment of the study is missing. This concerns studies that do not give sufficient experimental details and that are only listed in abstracts or secondary literature (books, reviews, etc.) or studies of which the documentation is not sufficient for assessment of reliability for one or more vital parameters.

### b. Reliability categories - nanomaterials

- *Reliable without restrictions:* All critical and important reliability criteria are fulfilled or partially fulfilled. The study is well designed, performed and documented. Nanomaterial properties and behaviour in the test system is extensively documented. The experiment has been carried out according to methods that are considered scientifically appropriate for ecotoxicity testing of nanomaterials and where the physicochemical properties of the nanomaterial are considered in the test design. If (when) specific nanomaterial guidance or guidelines exist, the use of these may be considered favourable.
- *Reliable with restrictions:* Most critical and important criteria are fulfilled or partially fulfilled. The study is generally well designed, performed and documented, but some minor flaws in the documentation or setup may be present. Nanomaterial properties and behaviour in the test system is well documented. The experimental design and test method are considered scientifically appropriate for ecotoxicity testing of nanomaterials but may contain some minor flaws in documentation or setup.
- *Not reliable:* Not all critical reliability criteria are fulfilled or partially fulfilled. This mainly concerns studies which have clear flaws in study design and study conduction, and/or where the experimental design and test method are considered not to be scientifically appropriate for ecotoxicity testing of nanomaterials.
- *Not assignable:* Information needed to make an assessment of one or more critical and important criteria is missing. This concerns studies or data from the literature which do not give sufficient experimental details, or reports where the documentation is not sufficient for assessment of reliability for one or more critical parameters.

**c. Relevance categories – all substances**

- *Relevant without restrictions:* The study is relevant for the purpose for which it is evaluated.
- *Relevant with restrictions:* The study has limited relevance for the purpose for which it is evaluated.
- *Not relevant:* The study is not relevant for the purpose for which it is evaluated.
- *Not assignable:* Studies that do not give sufficient details since the result is presented in abstracts or secondary literature (books, reviews, etc.) or studies of which the documentation is not sufficient for assessment of relevance for one or more vital parameters.

**References:**

Hartmann NB, Ågerstrand M, Holten Lützhof H-C, Baun A. 2017. "NanoCRED: A transparent framework to access the regulatory adequacy of ecotoxicity data for nanomaterial - Relevance and reliability revisited". NanoImpact, In press.

Kase R, Korkaric M, Werner I, Ågerstrand M. 2016. Criteria for Reporting and Evaluating ecotoxicity Data (CRED): comparison and perception of the Klimisch and CRED methods for evaluating reliability and relevance of ecotoxicity studies. Environmental Sciences Europe 28:7.

Moermond C, Kase R, Korkaric M, Ågerstrand M. 2015. "CRED - Criteria for Reporting and evaluating ecotoxicity Data." Environmental Toxicology and Chemistry 35: 1297-1309.

Molander L, Ågerstrand M, Beronius A, Hanberg A, Rudén C. 2014. "Science in Risk Assessment and Policy (SciRAP): An Online Resource for Evaluating and Reporting In Vivo (Eco) Toxicity Studies." Human and Ecological Risk Assessment. 21 (3), 753-762.

# Example from the SciRAP ecotoxicity tool

**Figure 1:** Print screen from the SciRAP ecotoxicity tool, reliability evaluation criteria. In this example, criteria 1 and 2 are faded since they have been removed in the evaluation.

Reliability
Reset form

---

1	Is a guideline method (e.g., OECD/ISO) or modified guideline used? (of minor importance for study reliability) <span style="color: green; font-weight: bold;">Guidance</span> ⓘ	Not determined ▾	OECD 201, not important for th
			<input type="button" value="Increase weight"/> <input type="button" value="Remove"/>
2	Is the test performed under GLP conditions? (of minor importance for study reliability) <span style="color: green; font-weight: bold;">Guidance</span> ⓘ	Not determined ▾	GLP lab, not important for the r
			<input type="button" value="Increase weight"/> <input type="button" value="Remove"/>
3	If applicable, are validity criteria fulfilled (e.g. control survival, growth)? <span style="color: green; font-weight: bold;">Guidance</span> ⓘ	Partially fulfilled ▾	Temperature not available
			<input type="button" value="Increase weight"/> <input type="button" value="Remove"/>
4	Are appropriate controls performed (e.g. solvent control, negative and positive control)? <span style="color: green; font-weight: bold;">Guidance</span> ⓘ	Fulfilled ▾	Solvent and negative control
			<input type="button" value="Increase weight"/> <input type="button" value="Remove"/>

---

17	Is a sufficient number of replicates used? Is a sufficient number of organisms per replicate used for all controls and test concentrations? <span style="color: green; font-weight: bold;">Guidance</span> ⓘ	Not fulfilled ▾	Unclear, is the guidance recomr
			<input type="button" value="Increase weight"/> <input type="button" value="Remove"/>
18	Are appropriate statistical methods used? <span style="color: green; font-weight: bold;">Guidance</span> ⓘ	Fulfilled ▾	Comment
			<input type="button" value="Increase weight"/> <input type="button" value="Remove"/>
19	Is a concentration-response curve observed? Is the response statistically significant? <span style="color: green; font-weight: bold;">Guidance</span> ⓘ	Fulfilled ▾	Comment
			<input type="button" value="Increase weight"/> <input type="button" value="Remove"/>
20	Are sufficient data available to check the calculation of endpoints and (if applicable) validity criteria (e.g., control data, concentration-response curves)? <span style="color: green; font-weight: bold;">Guidance</span> ⓘ	Not fulfilled ▾	Comment
			<input type="button" value="Increase weight"/> <input type="button" value="Remove"/>

Ågerstrand et al. 2012

Export to excel ⓘ

**Figure 2:** Print screen from the Excel sheet. Overview of the reliability evaluation results.

Selection	Total, accounting for weight	%, accounting for weight
Not determined	1	5,41%
Fulfilled	12	64,86%
Partially fulfilled	3,5	18,92%
Not fulfilled	2	10,81%

**Figure 3:** Print screen from the Excel sheet. Detailed report of the result from each criterion. In this example, criteria 1, 2 and 7 are faded since they have been removed in the evaluation.

Weight/ Removed	Test setup	Selection	Comment
Removed	Is a guideline method (e.g., OECD/ISO) or modified guideline used? (of minor importance for study reliability)	Not applicable	OECD 201, not important for the reliability of the study
Removed	Is the test performed under GLP conditions? (of minor importance for study reliability)	Not applicable	GLP lab, not important for the reliability of the study
1,5	If applicable, are validity criteria fulfilled (e.g. control survival, growth)?	Partially fulfilled	Temperature not available
1,5	Are appropriate controls performed (e.g. solvent control, negative and positive)	Fulfilled	Solvent and negative control
<b>Test compound</b>			
1	Is the test substance identified clearly with name or CAS-number? Are test results reported?	Fulfilled	No CAS-number
1,5	Is the purity of the test substance reported? Or, is the source of the test substance	Fulfilled	Source known, purity retrieved from them.
Removed	If a formulation is used or if impurities are present: Do other ingredients in the formulation exert an effect? Is the amount	Not applicable	No formulation
<b>Test organism</b>			
1	Are the organisms well described (e.g. scientific name, weight, length, growth, sex)?	Fulfilled	Scientific name, weight, length, growth and gender known
1	Are the test organisms from a trustworthy source and acclimatized to test conditions? Have the organisms not been pre-exposed	Fulfilled	
<b>Exposure conditions</b>			
1	Is the experimental system appropriate for the test substance, taking into account its characteristics?	Fulfilled	Static system
1	Is the experimental system appropriate for the test organism (e.g., choice of medium or test water, feeding, water characteristics, temperature, light/dark conditions, pH, etc.)?	Partially fulfilled	Information about temperature is missing.
1	Were exposure concentrations below the limit of water solubility (taking the use of a solvent into account)? If a solvent is used, is the solvent within the appropriate range?	Fulfilled	
1	Is a correct spacing between exposure	Fulfilled	
1	Is the exposure duration defined?	Fulfilled	According to the guideline
1	Are chemical analyses adequate to verify concentrations of the test substance over time?	Partially fulfilled	Some information about the method is missing.
1	Is the biomass loading of the organisms in the test system within the appropriate range?	Not determined	
<b>Statistical design and biological response</b>			
1	Is a sufficient number of replicates used? Is a sufficient number of organisms per concentration?	Not fulfilled	Unclear, is the guidance recommendation followed?
1	Are appropriate statistical methods used?	Fulfilled	
1	Is a concentration-response curve observed? Is the response statistically significant?	Fulfilled	
1	Are sufficient data available to check the calculation of endpoints and (if applicable) validity criteria (e.g., control data, etc.)?	Not fulfilled	

**Figure 4:** Print screen from the Excel sheet. The colour chart from the reliability evaluation.

