

Instructions for evaluating reliability and relevance of *in vivo* and *in vitro* toxicity studies using the SciRAP tool

Updated 23 February, 2018

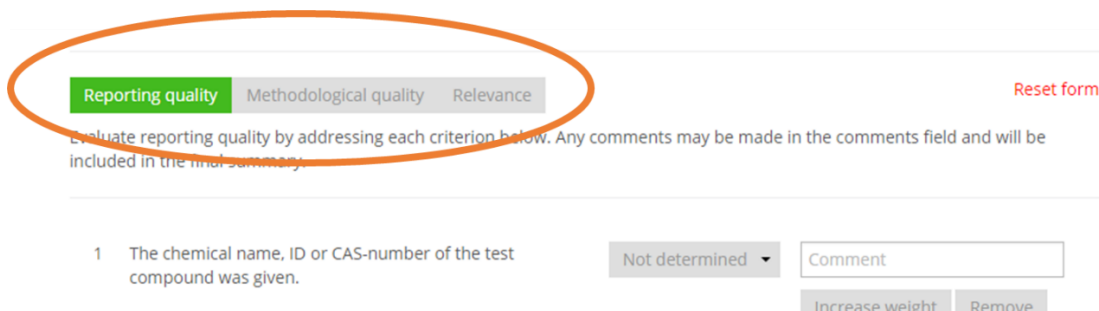
Contents

Introduction.....	1
Evaluation of Reliability.....	2
<i>Reporting quality</i>	2
<i>Methodological quality</i>	3
Evaluation of Relevance.....	3
Exporting the evaluation to Excel.....	4
Interpreting the output from the SciRAP tool.....	4

Introduction

To access the SciRAP tool for evaluating *in vivo* or *in vitro* studies, go to www.scirap.org. Select the tab for “*In vivo* toxicity studies” or “*In vitro* toxicity studies”, depending on the study you are evaluating, at the top of the page. Click on “Evaluate reliability & relevance”. You are now at the page with the evaluation tool for *in vivo* or *in vitro* studies, respectively. You may have to scroll down a bit to see it.

In the SciRAP method, reliability evaluation of *in vivo* and *in vitro* toxicity studies is conducted by evaluating “reporting quality” and “methodological quality” separately. The SciRAP criteria are thus organised under three different tabs in the tool, for “Reporting quality”, “Methodological quality” and “Relevance”. You may switch between these sets of criteria by selecting the corresponding tab (see figure below).



The screenshot shows the SciRAP tool interface. At the top, there are three tabs: "Reporting quality" (highlighted in green and circled in orange), "Methodological quality", and "Relevance". To the right of these tabs is a "Reset form" button. Below the tabs, there is a text prompt: "Evaluate reporting quality by addressing each criterion below. Any comments may be made in the comments field and will be included in the final summary." Below this, there is a list of criteria. The first criterion is "1 The chemical name, ID or CAS-number of the test compound was given." To the right of this criterion is a dropdown menu showing "Not determined" and a text input field labeled "Comment". Below the input field are two buttons: "Increase weight" and "Remove".

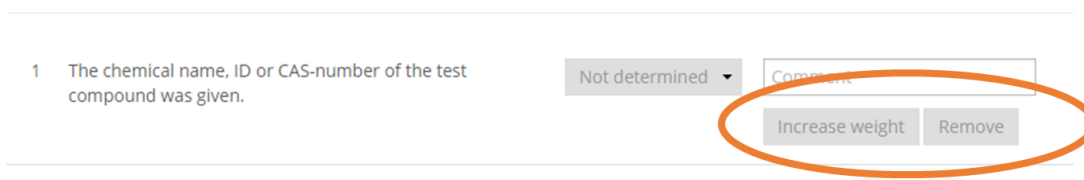
Although not a requirement, it is recommended to first evaluate reporting quality of the study before moving into the evaluation of methodological quality. Evaluation of reporting quality may identify

studies that have obvious deficiencies and/or are too poorly reported to allow for thorough evaluation of reliability and relevance. Further evaluation of clearly deficient studies may be avoided to save time and resources.

Removing or increasing the weight of criteria:

Individual criteria may be considered more or less critical in the specific case you are working on and the SciRAP tool includes a function to remove or increase the weight of individual criteria for reporting and methodological quality. **Note that this function is optional!**

The weight of a criterion can be increased by selecting the “Increase weight” button next to it. The weight of the criterion will then be increased from 1 to 1.5, which will be reflected in the output of the evaluation, i.e. the colour profile and the score (see section on “Interpreting the output from the SciRAP tool”). Criteria that are considered not applicable or not relevant in the specific case may be removed by choosing “Remove”. Removed criteria will not be included in the colour profile or calculation of the score. Motivations for weighing and removing criteria should be provided in the comments fields.



The screenshot shows a table with one row. The first column contains the number '1' and the text 'The chemical name, ID or CAS-number of the test compound was given.' The second column contains a dropdown menu with 'Not determined' selected. The third column contains a text input field with the placeholder text 'Comment'. Below the 'Comment' field are two buttons: 'Increase weight' and 'Remove'. These two buttons are circled in orange.

Please note that increasing the weight of or removing criteria will have an impact on the colour profile and the score! It is therefore important that the same criteria are weighed or removed in evaluations that are going to be compared to each other.

Note that your settings and selections for the criteria will be saved on your computer or device until you click “reset form”.

The evaluation conducted in the web based tool is exported to an excel file, which summarises the evaluation and presents colour profiles for reporting quality, methodological quality and relevance. The tool also calculates a score for reporting quality and methodological quality. **Certain considerations have to be made when interpreting the output of the SciRAP evaluation and especially the score.** These are discussed at the end of this document.

Evaluation of Reliability

Reporting quality

- To evaluate reporting quality of the study, select the “Reporting quality” tab. The criteria under this tab address different aspects of the reporting on the test compounds and controls, the test system, dosing and administration and data collection and analysis, as well as funding sources and competing interests.
- If needed, remove or increase the weight of specific criteria. Note: this is optional! Motivations for weighing and removing criteria should be provided in the comments fields.

- Evaluate how well each item (criterion) was reported in the study by selecting “fulfilled”, “partially fulfilled” or “not fulfilled” from the drop-down menu next to it. Any comments can be made in the comments field provided for each criterion.
- When all reporting quality criteria have been addressed, continue to evaluation of methodological quality by clicking the “Evaluate methodological quality” button at the bottom of the page, or by selecting the methodological quality tab at the top. Alternatively, if you are only interested in reporting quality you can export a summary of your evaluation by naming your evaluation and clicking the green “Export to excel” button at the bottom of the page (see “Exporting the evaluation to Excel”).

Methodological quality

- To evaluate methodological quality of the study, select the “Methodological quality” tab. Criteria under this tab address aspects of the study design, methods and conduct, which may influence the results of the study.
- If needed, remove or increase the weight of specific criteria. Note: this is optional! Motivations for weighing and removing criteria should be provided in the comments fields.
- Evaluate how well each methodological quality criterion was fulfilled in the study by selecting “fulfilled”, “partially fulfilled” or “not fulfilled” from the drop-down menu. *Note that guidance for evaluating each individual criterion is available by clicking the Guidance link next to it.* A criterion may be left as “not determined” if it cannot be evaluated due e.g. to insufficient reporting or lack of expertise. Any comments can be made in the comments field provided for each criterion.
- The evaluation of reporting quality and methodological quality completes the evaluation of study reliability in the SciRAP tool. Continue to evaluation of relevance by clicking the “Evaluate relevance” button at the bottom of the page, or by selecting the relevance tab at the top. Alternatively, if you are only evaluating study reliability you can export a summary of your evaluation by naming your evaluation and clicking the green “Export to excel” button at the bottom of the page (see “Exporting the evaluation to Excel”).

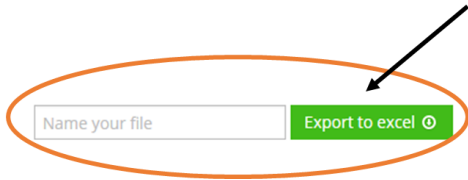
Evaluation of Relevance

The aim here is to evaluate to what extent the study is relevant for the assessment of human health hazards or risks.

- To evaluate relevance, select the relevance tab.
- Address each one of the relevance items. Each item should be judged as “directly relevant”, “indirectly relevant” or “not relevant” based on the context within which the evaluation is being conducted. That is, the specific hazard or risk assessment question under study. Guidance for evaluation is available for each aspect by clicking the question mark. Please note that all listed items do not have to be relevant for the study to serve as evidence in risk assessment.
- Relevance items cannot be removed or weighed up.

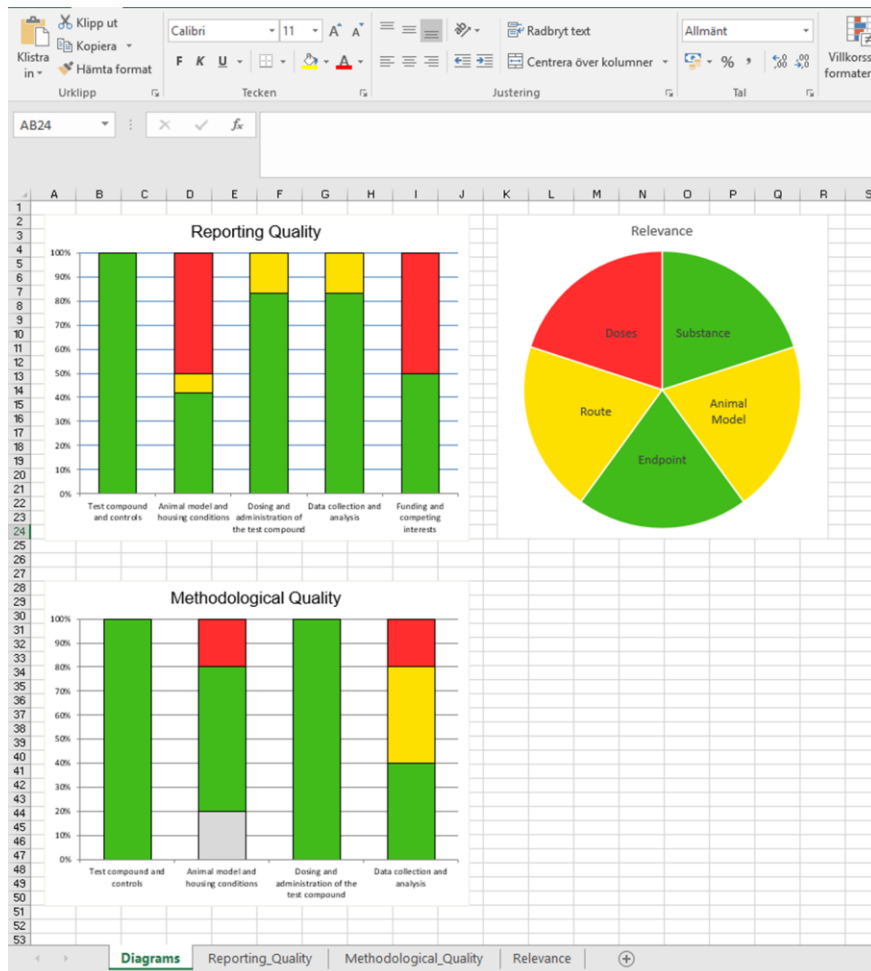
Exporting the evaluation to Excel

When the evaluation is completed, export a summary of your evaluation by naming it and clicking the green “Export to excel” button at the bottom of the page. The tool will generate an excel file, giving you a summary and a colour profile for reporting quality, methodological quality and relevance of the study, as well as numerical scores for both reporting quality and methodological quality. The excel file can be saved onto your computer.



Interpreting the output from the SciRAP tool

The first worksheet in the excel file gives an overall overview of the evaluation of reporting quality, methodological quality and relevance in the form of diagrams. This **colour profile** (see example below) of the study is specifically useful for illustrating where a study’s strengths and weaknesses lie.



The figure above shows an example of a colour profile generated with the SciRAP tool. Details about the evaluation can be found under the tabs for “reporting quality”, “methodological quality” and “relevance”, respectively.

In the colour profile, the evaluations of reporting quality and methodological quality are illustrated in bar charts, showing green for fulfilled criteria, yellow for partially fulfilled and red for criteria that were not fulfilled. Criteria that were “not determined” will be shown as grey. The bar charts take the weights of criteria into account and do not include criteria that have been removed. The evaluation of relevance is illustrated as a pie chart separate sections for each relevance item. Items evaluated as relevant are shown as green, indirectly relevant items are shown as yellow and if an item was evaluated as being not relevant for human hazard or risk assessment it is shown as red.

There are also separate worksheets that show the details of the evaluations of reporting quality, methodological quality and relevance, respectively. These show lists of the criteria and how each criterion has been evaluated. Any comments made for individual criteria will be listed. Removed criteria for reporting and methodological quality are greyed out in the list and the weight of criteria are shown in the first column.

The worksheets for reporting quality and methodological quality will also display a numerical score for reporting quality and methodological quality, respectively. The calculation of the scores can be simply explained as

$$\text{SciRAP score} = \frac{F + (PF * 0.5)}{T} * 100$$

where F is the number of fulfilled criteria multiplied by their individual weights, PF is the number of partially fulfilled criteria multiplied by their individual weights and T is the total number of criteria multiplied by their individual weights. Criteria that have been removed are excluded from the calculation. In other words, the score is the percentage of fulfilled and partially fulfilled criteria included in the evaluation, taking the weight of individual criteria into account, where partially fulfilled criteria contribute half the value as fulfilled criteria. The SciRAP score can have a value ranging from 0 (all criteria are judged as “not fulfilled”) to 100 (all criteria are judged as “fulfilled”).

Note that the SciRAP score is influenced by how criteria are weighed or removed! Consequently, scores can only be compared between evaluations where the same criteria have been removed or weighed up. Also, judging a criterion as “not fulfilled” will have the same impact on the score as leaving it as “not determined”. It is therefore advisable to leave as few criteria as possible as “not determined” and the user should take care to note the reason for leaving a criterion as “not determined”, i.e. if it is due to poor reporting in the study or lack of the user’s own expertise. Importantly, the score cannot be considered on its own but should be interpreted together with the colour profile when concluding on study reliability. The colour profile is crucial to identify where a study’s strengths and weaknesses lie and is more informative than the score for this purpose.

The output from the SciRAP evaluation can, for example, be used as basis for ranking studies according to their relative reliability or dividing studies into different categories of reliability, e.g. the Klimisch categories “reliable without restrictions”, “reliable with restrictions” and “not reliable”. However, ranking and categorization is done on a case-by-case basis and SciRAP does not dictate cut-off values or categorization based on SciRAP scores or colour profiles.

If you have any questions, please do not hesitate to contact us at anna.beronius@ki.se.