Instructions for evaluating the reliability and relevance of epidemiological studies using the SciRAPepi tool.

Introduction:

The SciRAPepi tool for evaluating epidemiological studies allows for evaluation of reliability and relevance. The evaluation often has to be endpoint-specific, meaning that the evaluation is carried out focusing on one of several endpoints investigated in the study. Separate evaluations may thus be necessary for different endpoints in one study. The evaluation may be conducted for either reliability, relevance, or both, depending on the purpose of the evaluation.

Download the excel files containing the assessment sheets available on the SciRAP website. Each excel file is tailored for specific epidemiological study design: cross-sectional, case-control, nested case-control, and cohort studies, as well as the file containing all criteria and items that are suitable for studies with no straightforward study design (**Fig. 1)**. Each sheet contain a brief introduction on using the SciRAPepi tool and pre-defined criteria/items to be evaluated in 2 sections for reliability and relevance.

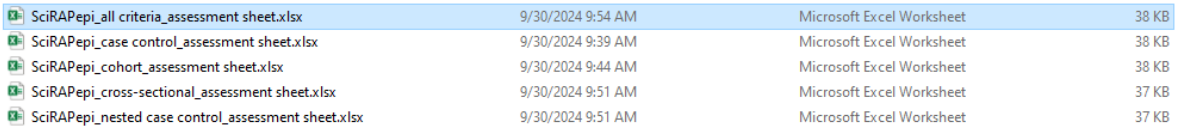


Fig. 1 *Separate SciRAP epi Excel files tailored for specific epidemiological study designs.*

The reliability section is divided in specific categories: Participants, Exposure measurement, Outcome measurement, Exposure and Outcome measurements, Data analysis, Ethics and competing interests, and Other (**Fig. 2**).

	RELIABILITY	SELECTION	COMMENT
	Participants		
1	The recruitment strategy and eligibility criteria for the participants were appropriate for the research question of the study.		
2	The sample size was appropriate for the statistical analysis and study design.		
3	The response rate of the potential participants was adequate.		
4	The comparison/control group was appropriate for the research question and study design.		
5	The follow-up of the participants was long enough to observe the outcome.		
6	The outcome was absent in participants at the beginning of the study.		
7	The loss to follow-up was ≤ 20% of the participants.		
8	The baseline characteristics of participants were described (demographic, social, health status) in this study, or a reference was added if characteristics were described previously.		
	Exposure measurement		
9	Reliable and sensitive methods were used for measuring the exposure (direct and indirect assessment methods).		
10	The same methods were used for measuring the exposure in all participants.		
11	Reliable (bio)matrices were used for measuring the exposure.		
12	Chemicals (reference material, substances/chemicals used in the pre-analytical and analytical phases) and laboratory equipment used for analysis were of appropriate quality and purity to reliably measure the (bio)markers of exposure.		
13	Reliable and specific (bio)markers were used for measuring the exposure.		
14	Reliable laboratory test procedures (e.g., laboratories with external quality assessment scheme certificate - EQUAS certificate, using internal and/or external quality control samples) were used when measuring the exposure.		
15	Appropriate percentage of samples was above the limit of detection (LOD) or limit of quantification (LOQ).		
16	Samples were treated appropriately and with a low risk of contamination during the pre-analytical and analytical phases.		
	Outcome measurement		
17	Reliable and sensitive methods were used for investigating the selected biomarker and/or outcome.		
18	The same methods were used to measure the biomarker/outcome in all participants.		
19	Reliable biomarkers were used for measuring the outcome.		
20	Reliable biomatrices were used for measuring the outcome.		
21	Chemicals (reference material, substances/chemicals used in the pre-analytical and analytical phases) and laboratory equipment used for analysis were of appropriate quality and purity.		
22	Reliable laboratory test procedures (e.g., laboratories with external quality assessment scheme certificate - EQUAS certificate, using internal and/or external quality control samples) were used when measuring the outcome.		
23	Appropriate number of samples was above the limit of detection (LOD) or limit of quantification (LOQ).		
24	Samples were treated appropriately and with a low risk of contamination during the pre-analytical and analytical phases.		
	Exposure & Outcome measurements		
25	The outcome assessors were blinded to the exposure assessment results, and exposure assessors were blinded to the outcome assessment.		
	Data analysis		
26	Concentrations of biomarkers were matrix adjusted (if needed).		
27	Appropriate data processing and statistical methods were used.		
28	Important confounders and effect modifiers were identified. These important confounders and effect modifiers were appropriately accounted for in the study design or analysis.		
29	Missing data were handled adequately in the dataset.		
30	If applicable, results (significant and not significant) of the statistical analyses were presented in the form of levels of significance, size of the effects (e.g., correlation coefficient, β-value), and accuracy (e.g., confidence intervals).		
31	Sensitivity analysis was performed to ensure the robustness of the results.		
	Ethics and competing interests		
32	The study was conducted with approval of Ethics Committee. Participants signed the informed consent at the beginning of the study.		
33	The funding sources for the study were stated and all competing interests were disclosed (or it was explicitly stated that the authors had no competing interests).		
	Other		
34	Other aspects of study design, performance or reporting that influence reliability.		

Fig. 2 *Categories of criteria in Reliability section of the SciRAPepi tool.*

Evaluation of the criteria:

When you evaluate the criteria/items, choose one of the options from the drop-down menu in the "SELECTION" column (fulfilled, partially fulfilled, or not fulfilled for reliability section; directly relevant, indirectly relevant, or not relevant in the relevance section, (**Fig. 3**). This drop-down menu is in almost every cell in the "SELECTION" column.

	RELIABILITY	SELECTION	COMMENT
	Participants		
1	The recruitment strategy and eligibility criteria for the participants were appropriate for the research question of the study.	fulfilled	
2	The sample size was appropriate for the statistical analysis and study design.	partially fulfilled	
3	The response rate of the potential participants was adequate.	not fulfilled	
4	The comparison/control group was appropriate for the research question and study design.	not reported	
5	The follow-up of the participants was long enough to observe the outcome.	REMOVE	
6	The outcome was absent in participants at the beginning of the study.	fulfilled	
7	The loss to follow-up was ≤ 20% of the participants.	partially fulfilled	
8	The baseline characteristics of participants were described (demographic, social, health status) in this study, or a reference was added if characteristics were described previously.	not fulfilled	
	Exposure measurement		
9	Reliable and sensitive methods were used for measuring the exposure (direct and indirect assessment methods).		
10	The same methods were used for measuring the exposure in all participants.		
11	Reliable (bio)matrices were used for measuring the exposure.		
12	Chemicals (reference material, substances/chemicals used in the pre-analytical and analytical phases) and laboratory equipment used for analysis were of appropriate quality and purity to reliably measure the (bio)markers of exposure.		

fulfilled
partially fulfilled
not fulfilled
not reported
REMOVE

Fig. 3 *Drop-down menu for the criteria in Reliability section of the SciRAPepi tool.*

Guidance for evaluating individual reliability criteria and relevance items is available by pointing to the criterion with the cursor (the criterion containing the guidance has a red right corner, **Fig. 4**).

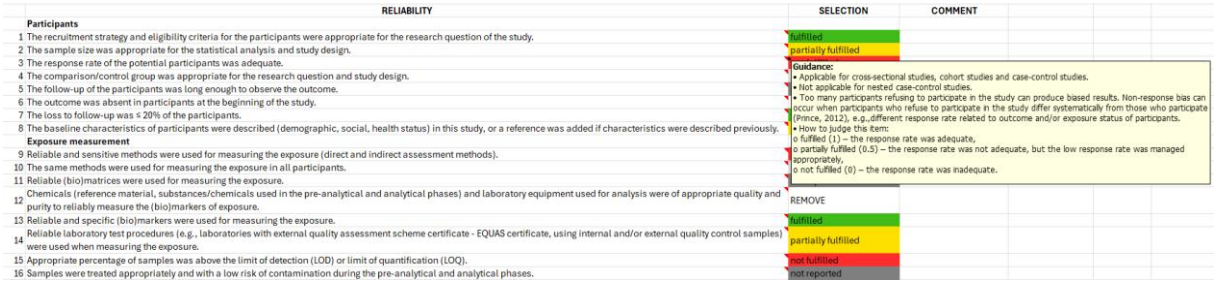


Fig. 4 *Guidance for evaluating each criterion in the SciRAPepi tool.*

Criterion no. 31 (or 34 in “cohort” and “all criteria” Excel files) provides space for free text comments on additional aspects that affect study reliability. These criteria do not contain the drop-down menu with options.

You may use the "COMMENT" column to write free text comments, for example explaining your evaluation of a specific criterion (**Fig. 5**).

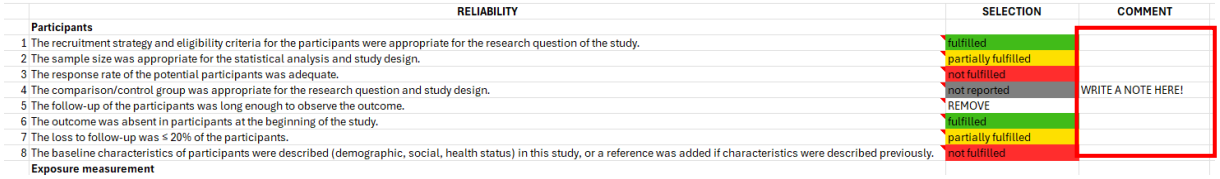


Fig. 5 *Writing a note in the "COMMENT" column.*

Judging criteria as “not reported”

If a criterion cannot be judged, you can select the option “not reported” in the drop-down menu (**Fig. 3**). This might be used when sufficient information is lacking to make a judgment regarding whether the criterion is fulfilled or not.

Removing criterion:

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 criteria for reliability. In that case, choose "REMOVE" in the drop-down menu of the "SELECTION" column instead of fulfilled, partially fulfilled, not fulfilled (**Fig. 3**). Removed criteria will not be included in the colour profile or % fulfilled criteria calculation. Motivations for removing criteria can be provided in the "COMMENT" column (**Fig. 5**).

NOTE: removing criteria will have an impact on the colour profile and the % fulfilled criteria. It is therefore important that the same criteria are removed in evaluations that are going to be compared to each other. Items in the Relevance section cannot be removed.

Interpreting the results of the SciRAPepi tool:

Results of the study assessment are shown right below the relevance section of the SciRAP tool in the form of % fulfilled criteria, as well as a colour profile.

	% FULFILLED CRITERIA
	RELIABILITY
Study overall	46.30
Participants	42.86
Exposure	41.67
Outcome	50.00
Exposure & Outcome	100.00
Analysis	30.00
Ethics & Competing interests	100.00

Fig. 6 *Table with % fulfilled criteria.*

Percent fulfilled criteria

The results show % fulfilled criteria of for the study overall, as well as for the specific criteria categories (**Fig. 6**).

- The % fulfilled criteria is calculated as follows:

$$\%\ fulfilled\ criteria = \frac{F + (PF * 0.5)}{T} * 100$$

where F is the number of fulfilled criteria, PF is the number of partially fulfilled criteria, and T is the total number of criteria. Partially fulfilled criteria contribute half the value as fulfilled criteria. Criteria that have been removed are excluded from the calculation.

The % fulfilled criteria can have a value ranging from 0 (all criteria are judged as "not fulfilled"/"not reported") to 100 (all criteria are judged as "fulfilled").

NOTE:

- selecting “not reported” for a criterion will have the same impact as “not fulfilled” on the % fulfilled value. The user should take care to note the reason for leaving a criterion as "not reported".
- removing criteria will have an impact on the % fulfilled criteria, as well as the colour profile. It is therefore important that the same criteria are removed in evaluations that are going to be compared to each other.
- importantly, the % fulfilled criteria 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 % fulfilled criteria for this purpose.

Colour profile

In the colour profile, the evaluations of reliability and relevance are illustrated in bar charts (**Fig. 7**), showing green for fulfilled criteria, yellow for partially fulfilled and red for criteria that were not fulfilled. Criteria that were "not reported" will be shown as grey. Relevance items evaluated as relevant are shown as green, indirectly relevant items are shown as yellow, and if the item was evaluated as being not relevant for the risk assessment or problem formulation, it is shown as red. The bar charts do not include criteria that have been removed.



Fig. 7 *The evaluations of reliability and relevance are illustrated in bar charts.*

Categorisation of reliability

A general methodology to reliability categorisation based on SciRAP evaluations for epidemiological studies is proposed in **Table 1**.

Table 1. *Rules for translating SciRAP evaluations into reliability categories (adapted from Holmer et al., 2024).*

Reliability category	**Rules for reliability categorisation**
Reliable without restriction	Percentage of fulfilled reliability criteria >80%*; all predefined key criteria are fulfilled; and there are no deficiencies in non-key criteria that may affect study reliability.
Reliable with restriction	Percentage of fulfilled reliability criteria >65%*; one or more predefined key criteria are partially fulfilled; or there are minor deficiencies in non-key criteria that may affect study reliability.
Not reliable	Percentage of fulfilled reliability criteria <65%*; one or more predefined key criteria are not fulfilled; or there are major deficiencies in key or non-key criteria that affect study reliability.
Not assignable	Reliability cannot be assigned due to insufficient reporting or because two or more predefined key criteria are not reported.

*the % used here is not prescriptive and should be decided on a case-by-case basis

It is recommended to apply an approach based on the use of predefined 'key criteria' considered critical for the reliability of the study in a specific assessment context. However, the selection of key criteria and exact categorisation principles should be decided on a case-by-case basis, depending on the research question, study type, endpoint, and intended use of the evidence. Examples are provided in e.g. Holmer et al., 2024; Wiklund & Beronius, 2022.

References:

Holmer ML, Zilliacus J, Draskau MK, Hlisníková H, Beronius A, Svingen T. 2024. Methodology for developing data-rich Key Event Relationships for Adverse Outcome Pathways exemplified by linking decreased androgen receptor activity with decreased anogenital distance. Reprod Toxicol. 128:108662. doi: 10.1016/j.reprotox.2024.108662. Epub ahead of print. PMID: 38986849.

Wiklund L and Beronius A. 2022. Systematic evaluation of the evidence for identification of endocrine disrupting properties of Bisphenol F. Toxicology. 476:153255. doi: https://doi.org/10.1016/j.tox.2022.153255

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