PHO MetaQAT Guide

Critical appraisal in public health

Critical appraisal is a necessary part of evidence-based practice and decision-making, allowing us to understand the strengths and weaknesses of evidence, and thus enabling us to make the best use of the evidence that is available. The use of the evidence-based model began in clinical medicine; adaptation is required for its use in public health practice, including changes in the approach to critical appraisal. The scope of appraisal should be expanded beyond the risk of bias, or internal validity, assessments found in clinical medical literature.[1] While the assessment of internal validity still forms the core of the appraisal process, it should be expanded to consider external validity and place both of these considerations within the context of the intended application of the evidence. We must also reduce our reliance on the traditional hierarchy of study design and instead consider what type of evidence is best for addressing the problem under consideration.[2-4]

The critical appraisal process involves making judgments about quality, drawing on knowledge of study design, topic area, and the intended use of the evidence. It is important to record these judgements in detail to complete an appraisal that is rigourous and transparent. While critical appraisal is important for all applications, it is particularly useful in public health where the evidence base is highly heterogeneous and may include non-traditional forms of evidence such as grey literature. Thorough documentation ensures accountability and transparency in the use of evidence.

PHO Meta-tool for quality appraisal

Many tools useful for appraising evidence in various health contexts have been identified in the existing body of literature.[1,5,6] These projects provided an excellent background on the issue and an overview of the tools available, but ultimately did not lead to a recommendation of an appraisal tool that can be consistently applied across multiple public health topics and projects. Following an approach developed by the Ontario Public Health Libraries Association,[7] the *Public Health Ontario Meta-tool for Quality Appraisal for Public Health Evidence (PHO MetaQAT)* was developed to provide a standard approach to critical appraisal for our organization. It is intended to increase the rigour and transparency of PHO products and processes by facilitating consistent completion and documentation of critical appraisal across the organization. The tool is designed to provide one standard process while also providing the flexibility needed for the variety of projects, topics, and staff experience at PHO.

We arrived at the concept of a meta-tool as means to create a tool that is both generic and specific, therefore providing broad utility while ensuring a high degree of rigour. The heterogeneity of public health evidence requires that a tool be generic and flexible to have utility for a wide range of projects across different topic areas. To ensure a high degree of rigour, it must also facilitate a detailed assessment of study-specific or item-specific (e.g. grey literature) factors. The meta-tool structure

provides one common appraisal process via a four-part framework, and also links to a suite of study-specific companion tools to provide the level of detail required for a rigorous examination of quality.

The meta-tool structure also allows the expansion of the appraisal process to include external validity and broader application while making use of existing, widely used tools to assess internal validity. The framework of the MetaQAT follows a four-part structure: relevancy; reliability; validity; and applicability. In the validity section, users are directed to the companion tools in order to provide a detailed assessment of internal validity if required. The extension of the appraisal process beyond internal validity led us to use the term "quality appraisal" when describing the meta-tool, rather than "critical appraisal". The generic framework was developed by cross-comparing existing tools and grouping related items to identify core quality appraisal concepts, which ensures a comprehensive assessment of quality. The companion tools were selected from existing critical appraisal tools in common use, with preference given for tools with documented validity and reliability.

> **Meta-tool: a tool that orients the user to the appropriate use of several appraisal tools and places them within a larger framework to guide their use**

The explicit consideration of relevancy, reliability, and applicability alongside validity is meant to make it clear that internal validity is one important consideration amongst others. The application of evidence in public health should draw on all parts of the framework together, and therefore the MetaQAT encourages the use of the hierarchy of study design to contribute to, rather than to dictate, the assessment of the evidence.

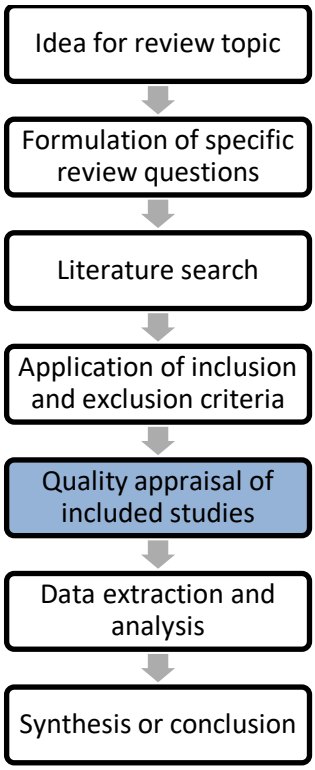
The tool is also designed to document all of these important details to provide transparency. Documentation in the MetaQAT takes the form of long-form written answers. This is the best format to capture the details of the strengths and weaknesses of the item and the contextual factors important to understanding how the evidence can best be used. There is no quantification in the MetaQAT; the result will not be a number that can be said to be "good" or "bad." While numeric scoring is common, it is not a valid approach to appraising quality.[1,8-11] Numerical summary scores mask important details: the same overall score may apply to both a study with one major flaw and to a study with several minor flaws.1

Companion tools which use numeric scales should be used for their content only – assessment should be made using the MetaQAT as a whole. A summary of the written answers will provide the information required to understand the strengths and weaknesses of the evidence, to make the appraisal process transparent, and to best understand the implications for applying the evidence to practice.

Purpose of the MetaQAT and its role in the research process

The MetaQAT form has been designed to help conduct and record the quality appraisal process, and integrate findings from diverse study designs. While quality appraisal concepts are useful to consider at

every stage of your work, the use of the MetaQAT form will take place roughly in the middle of your research process. An example process below shows where quality appraisal might occur in a research workflow:



Context: While some principles of quality appraisal might inform the development of a research question, and will contribute to the final conclusion of the research, the use of the MetaQAT form cannot be a substitute for the many other steps involved in a research project.

The public health research base is, in principle, more heterogeneous than that of clinical medicine. Public health deals with multi-modal interventions and issues in diverse settings. As such, appraising public health evidence via grading frameworks which favour quantitative designs like randomized controlled trials may unjustly penalize other types of research, such as quasi-experimental designs or qualitative stuides. Additionally, the practical needs of the public health field often require multiple types of study designs to be reviewed within a particular topic area thus requiring a framework that can adapt to multiple study designs. Many quality appraisal resources incorporate evidence grading concepts, but the MetaQAT is not an evidence grading tool.

Using the MetaQAT to critically appraise

The MetaQAT form is structured according to the four step framework: relevancy; reliability; validity; and applicability. An explanation of each of the four steps is provided below. Each step contains one or more main questions with a series of prompting questions that are meant to stimulate thinking around the domains. These supplementary questions are designed to help you consider different aspects of the

main question: do not feel you must answer each one in turn, or that no other items may be considered. You may find some prompting questions apply to your situation while others may not; you may also find it useful to customize the prompting questions for a particular application.

Next to each main question is a space provided to record your answer. As many of the questions are deliberately open-ended, the answer must specify how it applies and the particular strengths and weaknesses that relate to each question. The “yes”, “no”, “unclear”, and “n/a” checkboxes that accompany the answer are optional. It is the written answer that will provide a clear rationale for your appraisal, which results in transparency and can improve rigor.

Keep in mind that the written answers need to provide sufficient detail to draw from when summarizing the strengths and weaknesses of the evidence you have appraised. It is also useful to record anything about which you are unsure and want to investigate further or consult an expert, such as an unfamiliar statistical method you would like to flag for follow up with a statistician. In addition, the written answers are crucial to documenting your appraisal process and ensuring it is transparent. Another person should be able to understand your reasoning by reading the completed MetaQAT form.

While providing written answers is more time-consuming that using a form that relies on checkboxes, the appraisal will be more thorough and more transparent. Quality appraisal decisions are not simple - if evidence is to inform a decision in research, policy, or any other meaningful outcome, the significance of the appraisal process requires the investment of considerable time and effort. In many cases, that investment will pay off when writing up your research, as you’ll have your notes and written analysis already prepared, rather than just a checkbox to refer to. In addition, at the end of each section there is a summarizing box to allow for an overall judgment for each domain.

The appraisal form is also available in a spreadsheet format to facilitate appraising multiple items.

> **The form uses the term “study” to refer to the work being appraised and is meant to include all study types, including research published as grey literature, as syntheses, and any other type of evidence being considered.**

PHO MetaQAT framework: four steps to critical appraisal

1. ASSESSMENT OF RELEVANCY

Relevancy: whether you are assessing an individual article or conducting a large review, the first step is to determine if the item being appraised is sufficiently related to your topic or research question. In a structured systematic review, relevancy can also be thought of as the determination of whether the item meets the inclusion and exclusion criteria.

Timing

If you are completing a review, the assessment of relevancy will usually be done when screening your search results and not necessarily when you are conducting the full quality assessment. Note that screening may happen in two phases, with the first screening used to narrow down the search results to articles that need to be read in further detail to ensure they meet the inclusion and exclusion criteria, which would get screened at the second phase. The assessment of relevancy can also serve as a check of your inclusion and exclusion criteria and/or a chance to make refinements. The MetaQAT has been summarized into a screening aid called the quick QAT. However, if you are using the tool to appraise a single study, the relevancy section will be important in the context of your appraisal.

Assess

There are many ways to look at relevancy, depending on your situation. You may consider the study population and setting, and how similar and different they are to your situation. Consider the impact of these similarities and differences. Remember that you are unlikely to find an exact match – instead, focus on what you can learn from the evidence. Also consider the results of the study. Does the study examine outcomes you are interested in? Is the research design appropriate for your purpose? You may want to examine only quantitative or qualitative studies, or have a pre-set list of designs you will include in a review. Consider these issues and any others particular to your situation, and decide whether you will continue with the assessment of the item.

Document

Record the reasons why you consider the item relevant or not relevant in the answer space provided. Include relevant details like population, setting, etc. These contextual factors may be important in interpreting how to apply the evidence.

If you have found this article to be useful and wish to continue evaluating it, move on to part 2.

2. ASSESSMENT OF RELIABLITY

Reliability: complete reporting is necessary in order to conduct a thorough examination of quality, therefore the quality and transparency of reporting must be assessed. This section focuses on aspects which must be reported in sufficient detail for the evidence to be reproduced. In addition, lack of detail regarding the conduct of the study or report may be indicative of lower quality information.

Assess overall reporting

The first question is an initial assessment for completeness of reporting, to see whether there are any important omissions. Think of the first question as a quick scan for completeness of the item.

A research study should present a clear rationale for why the study was done. If you are appraising an item other than a journal article, you should still have some idea of the context, setting, and purpose.

The methods should be clearly and completely reported so that you can understand everything that was done. The results should be clear and complete, with all results included that you would expect to see given the methods described. Compare any data in tables and figures to the results reported and discussed in the text to ensure that they are consistent. The discussion should include a comparison of the study results to other research and discuss similarities and differences. Ensure the authors include some conclusions based on their results. Check to see whether conflict of interest information is provided.

Make notes about what is reported well and what is missing.

Assess detailed reporting of the methods and results

Look for sufficient detail in the description of the following:

- sample and sampling procedures, inclusion and exclusion criteria for participants, studies, or data
- sources of information, e.g., national survey data, hospital discharge abstracts, etc.
- study conduct e.g., interview protocols, survey design and distribution, outcome measurement, etc.
- analysis, whether using statistical or other methods

Overall, the methods and results should be presented clearly and completely so that you can understand exactly what was done and what was found. Consider whether you have sufficient information to replicate the study. Complete your written answer, including anything unclear or missing.

Ethics

Ethical considerations in public health are complex, and are treated in a distinct section to underscore their importance. Concern is placed primarily on the group level, with the need to also consider the interaction with individual level concerns.[12] Assessment should include the reporting of formal review procedures, and also be flexible to consider ethical implications more widely.

Look for a description of review by an ethics board, and of the informed consent process. Note that many older papers did not report this information. Consider whether there were additional concerns that should have been addressed.

Document

Ensure you have recorded all of the details mentioned in this section.

3. ASSESSMENT OF VALIDITY

In this section, you will consider in detail how the conduct of the study influences your confidence in the results. There are several research terms used to discuss the validity of a study:

Internal validity: the likelihood and magnitude of error or bias in a study. Things that contribute to high internal validity are things done to decrease the size and/or likelihood of errors, such as using a measure that has been previously validated.

External validity: the likelihood that the results could be generalized to a wider population, i.e., the results would be the same if studied in another setting. Things that contribute to high external validity ensure that the study setting, population, etc., are similar to the larger group of interest.

Bias: the result of something that decreases internal or external validity. It is a systematic error, and therefore influences the results in a particular way. This can limit the certainty of the conclusion and thus the application of the evidence may be limited. Biases are usually named according to the event that leads to the bias, e.g., social desirability bias – participants answer in a certain way due to social stigma associated with an opinion or behaviour.

To illustrate how these concepts work together, in the validity section of the MetaQAT we will examine the study conduct to determine potential biases and therefore our confidence in the results of the study (internal validity) and our confidence they would also apply elsewhere (external validity).

In order to answer the questions, you need to thoroughly understand how the study was conducted. If, in the previous section, you found the reporting to be inadequate then your ability to assess validity may be limited.

Focus on the methods

The aim of the first main question is to consider whether the methods were a good match for the research question. In the previous section, you considered whether you could understand the methods, now you will consider whether those methods are appropriate.

Consider appropriateness of the methodology as a whole, and within each component part. There should be clear congruence between the research aims and the design, the data, the analysis, and any other methodologies that were carried out. Was the design appropriate given the stated aims of the study? For example, ensure that if the authors claim a causal relationship, their design can support those claims. Were the data sources appropriate or were there more appropriate options? Consider whether the measures were the best choice for the concepts the authors were studying. Was the analysis the best choice considering the data and the research aims? Have the authors included all relevant factors in their analysis or are there potential confounding factors unaccounted for?

Record how and why the methods are appropriate or inappropriate.

Identify sources of bias

Consider how the different components of the study increase or decrease the likelihood of biases. In particular, consider the study design, participants included or excluded, measures used, sources of data, and quality assessment and selection of studies in the case of a systematic review. Are the authors' choices likely to influence the results in a particular way? Is the design comprehensive or are important factors missing that would provide a fuller understanding of the true situation?

The companion tools can provide design-specific questions to help in your assessment of bias.

Record potential sources of bias, and explain your concerns.

Compare the results and the conclusions

The results should support the authors' conclusions. Consider whether the conclusions are warranted or if they have been overstated. Would you make the same conclusions based on the results reported? Consider whether the results are meaningful. Do the authors discuss potential reasons for similarities and differences with other research? Are their explanations satisfactory?

Record your observations in your answer.

Assess overall confidence

Now that you have considered the strengths and weakness of the methodology and considered how the authors interpret their results, see how your thoughts compare. Your overall confidence in the results

can vary significantly based on the factors of reliability and validity that you have considered. Consider the main issues with the study that you have identified. Are they significant enough to limit the usefulness of the results? What are the caveats that should be kept in mind when looking at the results?

Document

Ensure you have recorded all of the relevant details for each of your answers, including strengths, weaknesses, and any questions to address.

4. ASSESSMENT OF APPLICABILITY

> **Applicability: in this section, consider how the evidence might be applied to public health practice. It is likely that most of the evidence will not be directly generalizable to your situation. Instead, consider more broadly what can be learned from the evidence and how you can apply that knowledge to decision making.**

Assess

Consider what you can learn from the study that could be applied in your situation. Is there something that could be applied to your program or policy? How can this evidence inform your decision making? Consider how the strengths and weaknesses you have identified in previous steps relate to your context.

Also consider any public health issues related to the evidence. Have the authors discussed the risks and benefits of an intervention? How does it relate to stakeholders concerns? How would it impact health equity in your context?

Document

Record the all important issues, both positive and negative, in your answer.

References

1. Heller RF, Verma A, Gemmell I, Harrison R, Hart J, Edwards R. Critical appraisal for public health: a new checklist. Public Health. 2008;122(1):92-8.
2. Harder T, Takla A, Rehfuess E, Sanchez-Vivar A, Matysiak-Klose D, Eckmanns T, et al. Evidence-based decision-making in infectious diseases epidemiology, prevention and control: matching research questions to study designs and quality appraisal tools. BMC Med Res Methodol. 2014;14:69,2288-14-69. Available from: http://www.biomedcentral.com/1471-2288/14/69;
3. Petticrew M, Roberts H. Evidence, hierarchies, and typologies: horses for courses. J Epidemiol Community Health. 2003;57(7):527-9. Available from: http://jech.bmj.com/content/57/7/527.long
4. Victora CG, Habicht JP, Bryce J. Evidence-based public health: moving beyond randomized trials. Am J Public Health. 2004;94(3):400-5. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1448265/
5. Bai A, Shukla V, Bak G, Wells G. Quality assessment tools project report. Ottawa, ON: Canadian Agency for Drugs and Technologies in Health; 2012. Available from: https://www.cadth.ca/media/pdf/QAT_final.pdf
6. European Centre for Disease Prevention and Control. Evidence-based methodologies for public health: how to assess the best available evidence when time is limited and there is a lack of sound evidence. Stockholm: European Centre for Disease Prevention and Control; 2011. Available from: http://ecdc.europa.eu/en/publications/Publications/1109_TER_evidence_based_methods_for_public_h ealth.pdf
7. Ontario Public Health Libraries Association (OPHLA). Critical appraisal of research evidence: a guide from the Ontario Public Health Libraries Association [Internet]. Toronto, ON: OPHLA; 2014 [cited 2015 May 5]. Available from: http://www.ophla.ca/pdf/CriticalAppraisalResearchEvidenceApril2014.pdf
8. Part 2: General methods for Cochrane reviews. 8.3.3: Quality scales and Cochrane reviews. In: Higgins JPT, Green S, editors. Cochrane Handbook for Systematic Reviews of Interventions. Version 5.1.0. London, UK: The Cochrane Collaboration; 2011. Available from: http://handbook.cochrane.org/
9. Herbison P, Hay-Smith J, Gillespie WJ. Adjustment of meta-analyses on the basis of quality scores should be abandoned. J Clin Epidemiol. 2006;59(12):1249-56.
10. Juni P, Witschi A, Bloch R, Egger M. The hazards of scoring the quality of clinical trials for metaanalysis. JAMA. 1999;282(11):1054-60.
11. Voss PH, Rehfuess EA. Quality appraisal in systematic reviews of public health interventions: an empirical study on the impact of choice of tool on meta-analysis. J Epidemiol Community Health. 2013;67(1):98-104.
12. McCarthy J. Ethics of public health. In: Wilson F, Mabhala A, editors. Key concepts in public health. Thousand Oaks, CA: SAGE Publications; 2009. p. 94-8.

Public Health Ontario | Santé publique Ontario

Public Health Ontario
661 University Avenue, Suite 1701
Toronto, Ontario
M5G 1M1
416.235.6556
communications@oahpp.ca
publichealthontario.ca

Ontario