

Critical Appraisal of Research Evidence

A Guide from the
Ontario Public Health Libraries Association

Copyright and Acknowledgements

© Ontario Public Health Libraries Association (OPHLA) 2008, updated 2014

Parts of this manual were built on content created by Susan J. Snelling for the Public Health Research, Education and Development (PHRED) program and presented at the Canadian Health Services Research Foundation's *Research Use Week (Northeastern Ontario): Tools, Strategies and Stories of Using Evidence in Rural and Remote Health Services Delivery and Policy Development* in May of 2007.

The OPHLA thanks Susan J. Snelling and PHRED for allowing this content to be adapted for use in this guide.

CONTENTS

Purpose of this guide	4
What is critical appraisal?	4
Why critically appraise research evidence?.....	4
Evidence-based medicine vs. evidence-based public health.....	5
When should you critically appraise?.....	5
How to critically appraise a research article.....	6
1. Is this article relevant to my issue and setting?	6
2. What are the author’s conclusions?	7
3. How confident can I be about the findings?	7
4. How can the results be applied to public health practice?	9
Quick reference: key questions	10
Key questions for critical appraisal of public health research	10
Critical appraisal form for public health research	11
Online critical appraisal resources.....	13
Recommended reading	14
References	15

PURPOSE OF THIS GUIDE

The purpose of this guide is to provide a brief overview of the critical appraisal process. Assessing the validity of research studies can be a complex and time-consuming undertaking. If you are conducting a lengthy evaluation, you may wish to consult more exhaustive critical appraisal resources (a list of suggested further reading has been appended to this guide). Participation in the *Skills Enhancement for Public Health* program offered by the Public Health Agency of Canada is recommended prior to attempting in-depth critical appraisal.

WHAT IS CRITICAL APPRAISAL?

Critical appraisal is "the process of carefully and systematically examining research to judge its trustworthiness, and its value and relevance in a particular context."¹ Critical appraisal is an essential step in the process of putting research into practice. Asking questions about an article's research methodology, scrutinizing its data collection and analysis methods, and evaluating how its findings are presented will help you to determine whether that article's conclusions should influence practical decision-making.

WHY CRITICALLY APPRAISE RESEARCH EVIDENCE?

It is crucial to critically evaluate research evidence in order to facilitate evidence-based practice, which is the use of the best evidence available to guide decision making and program design. The term "best evidence" emphasizes the fact that it is the quality and not the quantity of evidence that is of primary significance. Critical appraisal allows you to distinguish the best available evidence from within a large body of research.

Using a best-evidence approach allows you to:

- Retrieve reliable, up-to-date information about which interventions do and do not work for a particular public health topic;
- Control the amount of literature that you will need to analyze; and
- Feel confident that public health decision making is based on the "best of the best" information available on a particular topic.

EVIDENCE-BASED MEDICINE VS. EVIDENCE-BASED PUBLIC HEALTH

Fundamental differences between the fields of medicine and public health demand unique approaches to evidence-based practice for each discipline. Evidence-based public health is defined as the "process of integrating science-based interventions with community preferences to improve the health of populations."² The chart below outlines the major differences between evidence-based medicine and evidence-based public health.

		Evidence-Based Medicine	Evidence-Based Public Health
Definition		The process of systematically finding, appraising, and using contemporaneous research findings as the basis for clinical decisions.	The process of systematically finding, appraising, and using contemporaneous clinical and community research findings as the basis for decisions in public health.
Steps in the Process	1	Formulating a clear question from a patient's problem	Formulating a clear question from a public health problem
	2	Searching the literature	Searching the literature
	3	Appraising the evidence	Appraising the evidence
	4	Selecting the best evidence for clinical decision	Selecting the best evidence for a public health decision
	5	Linking evidence with clinical experience, knowledge, practice, and the patient's values and preferences	Linking evidence with public health experience, knowledge, practice, and the community's values and preferences
	6	Implementing findings in clinical practice	Implementing findings in public health practice and programs
	7	Evaluating results	Evaluating results
Goal		The best possible management of health and disease in individual patient(s)	The best possible management of health and disease and their determinants at the community level

Source: Jenicek & Stachenko, 2003.³ Reproduced with permission.

WHEN SHOULD YOU CRITICALLY APPRAISE?

Just a few of the instances in which it is important to use critical appraisal include:

- Conducting literature reviews for grant proposals;
- Evaluating the effectiveness, costs, and benefits of health programs;
- Establishing new health programs;
- Implementing policies; and
- Public health decision making, especially at the senior management level.

HOW TO CRITICALLY APPRAISE A RESEARCH ARTICLE

The way in which you critique research evidence will differ slightly according to the type of study you are appraising (e.g. randomized control trials (RCTs), systematic reviews, observational studies, meta-analysis, etc.).

Systematic reviews are a reliable source of evidence due to the fact that they appraise and summarize numerous primary research studies. However, they are not available on every public health topic. If a systematic review related to your topic has not been conducted, you will have to look at other types of studies in order to find the best available evidence.

Once you have conducted a literature search and obtained full text articles, you can begin the critical appraisal process. Consider the following questions to assess the quality of the article.

1. IS THIS ARTICLE RELEVANT TO MY ISSUE AND SETTING?

1.1. READ THE ABSTRACT

Use the information found in the abstract to answer the questions below.

- Are your issues discussed there?
- What are the main findings of the research?
- Do you want to know more after reading the abstract?
- Was the research done in a similar setting to yours?
- Does it address a related question? (Even research that covers your issue indirectly can be useful.)
- Are there reasons to doubt the findings without reading the whole article?

You may conclude that the study is not reliable merely by reading the abstract. If this is the case, move on to another article.

1.2. READ THE INTRODUCTION AND DISCUSSION SECTIONS

To further assess the relevance of a study, you will look to those sections of the research article which describe in more detail the objectives and context of that study. The Introduction and Discussion sections will help you to identify the key concepts, goals, subjects, and themes of the research.

Focus on the time, place and circumstances of the population and setting studied. How similar or different is the study population or setting to yours? Is a difference likely to matter for the issue at hand?

1.3. CONSULT THE METHODS SECTION

The Methods section will give you a step-by-step description of exactly how the study was carried out. In this section you should take note of:

- Where the study was done (on site, at home, administrative setting, etc.);
- From whom the data was collected (primary from staff, patients, families; or secondary from databases); and

- How the data was collected (interviews, focus groups, questionnaires, surveys, observations, etc.).

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

2. WHAT ARE THE AUTHOR'S CONCLUSIONS?

2.1. COMPARE THE ABSTRACT TO THE DISCUSSION SECTION

The discussion section is more detailed and precise than the abstract, and will explain the limitations of the research and possible implications which are not mentioned in the abstract.

2.2. COMPARE THE RAW DATA CONTAINED IN TABLES WITH THE RESULTS ANALYZED IN THE DISCUSSION AND CONCLUSIONS SECTIONS

Are the results reported in the conclusions consistent with what is reported in the tables? Is the interpretation consistent with what the actual findings were? They should be consistent, but aren't always.

2.3. HOW WELL ARE THE RESULTS RELATED TO OTHER RESEARCH ON THE SAME TOPIC?

In the Discussion or Conclusions section, is there a review of how these results compare or contrast with prior research? If this report found something different from previous research, then it's important to progress to question 3 on appraising the reliability of the findings.

3. HOW CONFIDENT CAN I BE ABOUT THE FINDINGS?

Peer-reviewed journals provide some measure of 'quality control' for formally published research. Peer-reviewed journals require submissions to pass through rigorous review by experts in the field before they are published. You can tell if a journal is peer reviewed by looking inside the front cover or in the submission requirements for the journal. However, even peer-reviewed resources vary in quality.⁴

To determine if a study's findings are trustworthy, you will review the Methods section. There are five factors that influence the reliability of research results:

- Completeness of the model that is analyzed (if there is one)
- Quality and relevance of the measures used and their relationship to the model
- Quality of the data
- Ability to control for differences between groups being compared
- Appropriateness of the statistical methods given the nature of the data generated

3.1. HOW COMPLETE IS THE MODEL?

A model is a description of the relationship between a dependent variable and the outcome with which it is believed to be associated. A model may also specify how the dependent and independent variables are conceptually related to one another. The following questions will help you to evaluate the completeness of a model:

- Are all the relevant factors included in the research?
 - How complete/relevant is the theory?
 - Are important factors or variables left out of the theory?
 - Are important theoretical factors accounted for in the analysis?
 - Does the model explain the “facts” as you currently understand them? If not, re-examine both the “facts” and your understanding of them.
- How important are the variables that may have been left out?
 - Does the study take, for example, socioeconomic or other variables into account in a reasonable way?
 - Does the study consider special contextual events, study location or patient characteristics that might influence the results?
 - Are omitted variables correlated with important policy or program variables? How would this affect the results?

3.2. HOW GOOD ARE THE MEASURES?

- Do the measures accurately reflect what the researcher was trying to measure (validity)?
- How clear and appropriate are these measures? (Too broad? Too narrow? Ambiguous?)
- Are they actual measures or proxy measures?
- Are the measures well established in either prior research or through pilot testing by the researcher, or are they ad hoc?

3.3. HOW GOOD IS THE DATA?

Measures (and the conclusions drawn from them) are only as good as the data used to construct them. Lots of missing data or a poor response rate can limit the reliability of the findings.

3.4. DOES THE STUDY ADEQUATELY CONTROL FOR DIFFERENCES BETWEEN THE GROUPS BEING COMPARED?

Many studies at least implicitly involve comparison across groups (control vs. intervention, age groups, genders, etc) on defined outcomes. The study should minimize (i.e. control for) all other differences between the groups being compared other than the variable being measured by the study. This is what randomization tries to achieve in clinical trials.

If this control is impossible (which is often the case in health services research due to ethical concerns), the study should adequately control for differences by other methods (e.g. matching).

Of particular importance is the selection of subjects. This is complicated because subjects assigned to the intervention group may be different from those assigned to the control group in ways that can't be directly measured. However, there are a wide range of methods to control for selection, such as the use of statistical adjustments (e.g. stratifying by age or illness level). Ideally, a research article will acknowledge the need to control for selection bias and will describe the researcher's method for doing so.

- How similar or different are the groups being compared?
 - If the groups are different, how would you expect the differences to influence the outcome being studied?
 - How have the researchers tried to control for these differences?
- Is there a risk of selection bias?
 - How have the researchers addressed this?
 - Note there are also a host of other biases (some quite subtle) that have an impact on research.

3.5. ARE THE STATISTICAL METHODS APPROPRIATE?

You may not be a statistician or methodologist, but the following questions can help sort out this issue:

- What was the reason the research was done?
 - Was the method a good match for the purpose? If the purpose is to understand the patient experience, a survey or qualitative study may be the best design; if the purpose is to determine which treatment has better outcomes, a controlled trial is a better match.
- Is the sample large enough to produce significant results? Is the effect size clinically or operationally relevant?
 - Relevant differences may not be "statistically significant" in a small sample. Is this discussed in the case of negative findings? If results are not statistically significant, were the study groups being compared too small to be able to detect an effect?
- Is there a discussion of how the methods relate to those used in other studies?
 - Sometimes authors will explicitly state their reasons for choosing the research methods that they have used; this may indicate to you whether they have addressed study design flaws identified in previous research.

4. HOW CAN THE RESULTS BE APPLIED TO PUBLIC HEALTH PRACTICE?

The application of research findings to community interventions is an integral part of public health practice. It is important to consider how the results of a study can realistically be integrated into local practice. The following questions will guide your assessment of the applicability of a study's results:

- How can the results be interpreted and applied to public health? Can I apply it to my program/policy?
- Were all important public health outcomes considered?
- Are the benefits worth the costs and potential risks?

QUICK REFERENCE: KEY QUESTIONS

The following summary has been formatted for quick reference. Print this page and keep it close at hand when conducting critical appraisal of your public health research results.

KEY QUESTIONS FOR CRITICAL APPRAISAL OF PUBLIC HEALTH RESEARCH

IS THE ARTICLE RELEVANT TO YOUR TOPIC?

- Are the issues discussed in the abstract of interest to you?
- Was the research done in a similar setting to yours?

WHAT ARE THE RESULTS?

- Were the results similar from study to study?
- What are the overall results of the study?
- How precise are the results?
- Can a causal association be inferred from the available data?

ARE THE RESULTS VALID?

- Did the review explicitly address the public health question?
- Was the search for relevant studies detailed and exhaustive?
- Were the primary studies of high methodological quality?

CAN THE RESULTS BE APPLIED TO PUBLIC HEALTH PRACTICE?

- Can I apply the results to my program/policy?
- Are all important public health outcomes considered?
- Are the benefits worth the costs and potential risks?

Source: Partners in Information Access for the Public Health Workforce., 2013.⁵ Adapted with permission.

CRITICAL APPRAISAL FORM FOR PUBLIC HEALTH RESEARCH

Print one form for each article that you will be appraising. Copy the citation for the article that you will be appraising into the grey box below. Record your answer to each question and keep track of your comments. Keep this form as a record of your appraisal for future reference.

Citation:

<p>Is this article relevant to my issue and setting?</p> <p>Does the study address a topic related to my research question?</p> <p>Was the research conducted in a setting similar to mine?</p>	<input type="checkbox"/> YES <input type="checkbox"/> NO	<p>Comments:</p>
<p>Are the results presented objectively?</p> <p>Are the results from all included studies clearly displayed?</p> <p>Are the results similar to those found by other studies on the same topic?</p>	<input type="checkbox"/> YES <input type="checkbox"/> NO	<p>Comments:</p>
<p>Are the author's conclusions justified?</p> <p>Is there a conclusive result?</p> <p>Are there any numerical outcomes?</p> <p>Are the results reported in the data tables consistent with those described in the Discussion and Conclusions sections?</p> <p>Are potential discrepancies discussed?</p>	<input type="checkbox"/> YES <input type="checkbox"/> NO	<p>Comments:</p>

<p>Is the research methodology clearly described and free of bias?</p> <p><i>For review articles:</i></p> <p>Is there a list of the specific bibliographic databases that were searched?</p> <p>Are the search terms listed?</p> <p>Are informal information sources included (grey literature, expert opinion, etc.)?</p> <p>Are non-English language articles included?</p> <p><i>For primary studies:</i></p> <p>Are the results precise (is there a confidence interval)?</p> <p>Are the statistical methods appropriate? Can the study be reproduced?</p>	<p><input type="checkbox"/> YES</p> <p><input type="checkbox"/> NO</p>	<p>Comments:</p>
<p>Can I be confident about findings?</p> <p>Does the study have a clearly stated objective and focus on a clearly defined issue?</p> <p>Does it describe the population studied, the intervention given, and the outcomes?</p> <p>Is the model that is being analyzed complete?</p> <p>Are the data valid and of good quality?</p> <p>Were the included studies quality assessed? What measures did they use?</p> <p>Is there a control group?</p>	<p><input type="checkbox"/> YES</p> <p><input type="checkbox"/> NO</p>	<p>Comments:</p>
<p>Should I apply the results to local public health practice?</p> <p>Can the results be interpreted and applied within the scope public health practice?</p> <p>Are the benefits worth the potential harms and costs?</p> <p>Are all important public health outcomes considered?</p>	<p><input type="checkbox"/> YES</p> <p><input type="checkbox"/> NO</p>	<p>Comments:</p>

AGREE II | AGREE Research Trust

AGREE II is the new (2010) international tool to assess the quality and reporting of practice guidelines. The original AGREE Instrument was published in 2003 by a group of international guideline developers and researchers, the AGREE Collaboration. The assessment includes judgments about the methods used for developing the guidelines, the components of the final recommendations, and the factors that are linked to their uptake

Available from: http://www.agreetrust.org/wp-content/uploads/2013/10/AGREE-II-Users-Manual-and-23-item-Instrument_2009_UPDATE_2013.pdf

CASP Critical Appraisal Tools | Critical Appraisal Skills Program International

The CASP International Network (CASPin) is a non-profit making organisation for people promoting skills in finding, critically appraising and acting on the results of research papers. Appraisal tools are available for systematic reviews, randomised controlled trials, qualitative research studies, economic evaluation studies, cohort studies, case control studies, diagnostic test studies, and clinical prediction rules.

Available from: <http://www.caspinternational.org/?o=1012>

Finding and Appraising the Evidence Video Course | HealthKnowledge/Public Health Action Support Team

This course has been written for those who wish to improve their skills in critical appraisal. The course starts from the basics and does not assume prior knowledge or training in this area. The course has been structured so that busy health care professionals, and others, can undertake learning in small chunks whenever they have a spare moment.

Available from: <http://www.healthknowledge.org.uk/interactive-learning/finding-and-appraising-the-evidence>

The IDM Manual: Evidence Framework | University of Toronto

The IDM Manual is a guide to the IDM (Interactive Domain Model) Best Practices Approach to Better Health. It is written from the perspective of health promotion practitioners. The manual outlines considerations about information relevance and quality, including related worksheets. The framework also includes critical appraisal logic models for identifying and assessing evidence from individual research/evaluation studies and review articles.

Available from: http://www.utoronto.ca/chp/download/IDMmanual/IDM_evidence_dist05.pdf

Rapid Evidence Assessment (REA) Toolkit | Government Social Research Service

This toolkit from the UK's Government Social Research Service (GSR) has been designed as a web-based resource to enable researchers to carry out or commission Rapid Evidence Assessments, also known as REAs. It contains detailed guidance on each stage of an REA from deciding on whether it is the right method to use to communicating the findings. There are a range of templates and sources in the Toolkit that will support the successful completion of an REA.

Available from: <http://www.civilservice.gov.uk/networks/gsr/resources-and-guidance/rapid-evidence-assessment>

Scottish Intercollegiate Guidelines Network (SIGN) has developed a number of critical appraisal tools in support of their creation of evidence-based guidelines. These appraisal tools used by SIGN to conduct systematic reviews have been made available for public use on their website, including methodology evaluation checklists for each of the following types of studies: systematic reviews/meta analyses, RCTs, cohort studies, case-control studies, diagnostic studies, and economic evaluations.

Available from: <http://www.sign.ac.uk/methodology/checklists.html>

RECOMMENDED READING

Benos DJ, Basharia E, Chaves J, Gaggar A, Kapoor N, LaFrance M. et al. The ups and downs of peer review. *Adv Physiol Educ* 2007;31:145-152.

Brouwers MC, Kho ME, Browman GP, Burgers JS, Cluzeau F, Feder G, Fervers B, Graham ID, Hanna SE, Makarski J; AGREE Next Steps Consortium. Development of the AGREE II, part 2: assessment of validity of items and tools to support application. *CMAJ*. 2010;182(10):E472-8.

Burls A. What is critical appraisal? 2nd ed. Oxford, UK: University of Oxford; 2009. (What is...? series). Available from: http://www.medicine.ox.ac.uk/bandolier/painres/download/whatis/what_is_critical_appraisal.pdf

Ciliska D, Thomas H, Buffet C. An introduction to evidence-informed public health and a compendium of critical appraisal tools for public health practice (revised). Hamilton, ON: National Collaborating Centre for Methods and Tools; 2012. Available from: http://www.nccmt.ca/pubs/IntroEIPH_compendiumENG_web.pdf

Crombie IK. The pocket guide to critical appraisal: a handbook for health care professionals. London: BMJ Publishing Group; 1996.

Greenhalgh T. How to read a paper: the basics of evidence-based medicine. 5th ed. Chichester, UK: Wiley-Blackwell; 2014. Available from: <http://ca.wiley.com/WileyCDA/WileyTitle/productCd-EHEP003117.html>

Gugiu PC, Gugiu MR. A critical appraisal of standard guidelines for grading levels of evidence. *Eval Health Prof.* 2010;33(3):233-55.

Guyatt G, Rennie D, Meade MO, Cook DJ, editors. Users' guides to the medical literature: a manual for evidence-based clinical practice. 2nd ed. 2008. Available from: <http://jamaevidence.com/resource/520>

Hannes K, Lockwood C, Pearson A. A comparative analysis of three online appraisal instruments' ability to assess validity in qualitative research. *Qual Health Res.* 2010;20(12):1736-43.

IDM Best Practices [homepage on the Internet]. Kahan B, Goodstadt M, ed. Regina, SK: IDM Best Practices; n.d. References to assess sources of evidence for health promotion, public health and population health; n.d. [cited 2012 Oct 26]. Available from: <http://www.idmbestpractices.ca/idm.php?content=resources-asseesev>

Jeanfreau SG, Jack L Jr. Appraising qualitative research in health education: guidelines for public health educators. *Health Promot Pract.* 2010;11(5):612-7. Available at: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3012622/pdf/nihms257124.pdf>

Kittle JW, Parker R. *How to read a scientific paper*. Tucson, AZ: University of Arizona; 2006. Available from: <http://www.biochem.arizona.edu/classes/bioc568/papers.htm>

Loney P, Chambers L, Bennett K, Roberts J, Stratford P. Critical appraisal of the health research literature: prevalence or incidence of a health problem. *Chronic Dis Can* 2000;19(4):1-13. Available from: http://www.collectionscanada.gc.ca/webarchives/20071212045422/http://www.phac-aspc.gc.ca/publicat/cdic-mcc/19-4/e_e.html

Lumley J, Daly J. In praise of critical appraisal. *Aust N Z J Public Health*, 2006;30(4):303. Available from: http://www.blackwellpublishing.com/pdf/anzjph/editorials_on_methods/praise_critical_appraisal_aug_2006.pdf

Ontario Public Health Libraries Association. *Literature searching for evidence: a reading list*. Toronto, ON: OPHLA; 2007. Available from: <http://www.ophla.ca/pdf/Literature%20Searching%20for%20Evidence%20Reading%20List.pdf>

Rychetnik L, Frommer M. *A schema for evaluating evidence on public health interventions*. Sydney, Australia: University of Sydney; 2002. Available from: <http://www.nphp.gov.au/publications/phpractice/schemaV4.pdf>

Stetler CB, Damschroder LJ, Helfrich CD, Hagedorn HJ. A guide for applying a revised version of the PARIHS framework for implementation. *Implement Sci*. 2011;6:99. Available at: <http://www.implementationscience.com/content/6/1/99>

Vlayen J, Aertgeerts B, Hannes K, Sermeus W, Ramaekers D. A systematic review of appraisal tools for clinical practice guidelines: multiple similarities and one common deficit. *Int J Qual Health Care*. 2005;17(3):235-42. Available at: <http://intqhc.oxfordjournals.org/content/17/3/235.full.pdf>

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.

REFERENCES

1. Burls A. *What is critical appraisal?* 2nd ed. Oxford, UK: University of Oxford; 2009. (What is...? series). Available from: http://www.medicine.ox.ac.uk/bandolier/painres/download/whatis/what_is_critical_appraisal.pdf
2. Brownson RC, Baker EA, Leet TL, Gillespie KN, True WR, editors. *Evidence-based public health*. 2nd ed. Oxford, UK: Oxford University Press; 2010. Chapter 1, The need for evidence-based public health; p. 5.
3. Jenicek M, Stachenko S. Evidence-based public health, community medicine, preventive care. *Med Sci Monit*. 2003 Feb;9(2):SR1-7. p. SR2.
4. Benos DJ, Bashari E, Chaves JM, Gaggari A, Kapoor N, LaFrance M, et al. The ups and downs of peer review. *Adv Physiol Educ*. 2007;31(2):145-52.
5. Partners in Information Access for the Public Health Workforce. *Public health information and data tutorial*. Rockville Pike, Bethesda: National Library of Medicine; [updated 2013 Jan 4]. Evidence-based public health: key concepts: steps in searching and evaluating the literature. Available from: <http://phpartners.org/tutorial/04-ebph/2-keyConcepts/4.2.5.html>