



ELSEVIER

Available online at www.sciencedirect.com

Public Health

journal homepage: www.elsevier.com/puhe

Original Research

The development and validation of a meta-tool for quality appraisal of public health evidence: Meta Quality Appraisal Tool (MetaQAT)



L. Rosella ^{a,b,c,*}, C. Bowman ^b, B. Pach ^b, S. Morgan ^b, T. Fitzpatrick ^b,
V. Goel ^{a,b,c,d}

^a Dalla Lana School of Public Health, University of Toronto, Health Sciences Bldg, 6th Floor, 155 College St., Toronto, Ontario M5T 3M7, Canada

^b Public Health Ontario, 480 University Avenue, Suite 300, Toronto, Ontario M5G 1V2, Canada

^c Institute for Clinical Evaluative Sciences, Canada

^d Institute for Health Policy Management and Evaluation, University of Toronto, Canada

ARTICLE INFO

Article history:

Received 14 February 2015

Received in revised form

30 September 2015

Accepted 29 October 2015

Available online 16 March 2016

Keywords:

Quality appraisal

Synthesis

Critical appraisal

Public health evidence

ABSTRACT

Objectives: Most quality appraisal tools were developed for clinical medicine and tend to be study-specific with a strong emphasis on risk of bias. In order to be more relevant to public health, an appropriate quality appraisal tool needs to be less reliant on the evidence hierarchy and consider practice applicability. Given the broad range of study designs used in public health, the objective of this study was to develop and validate a meta-tool that combines public health-focused principles of appraisal coupled with a set of design-specific companion tools.

Study design: Several design methods were used to develop and validate the tool including literature review, synthesis, and validation with a reference standard.

Methods: A search of critical appraisal tools relevant to public health was conducted; core concepts were collated. The resulting framework was piloted during three feedback sessions with public health practitioners. Following subsequent revisions, the final meta-tool, the Meta Quality Appraisal Tool (MetaQAT), was then validated through a content analysis of appraisals conducted by two groups of experienced public health researchers (MetaQAT vs generic appraisal form).

Results: The MetaQAT framework consists of four domains: relevancy, reliability, validity, and applicability. In addition, a companion tool was assembled from existing critical appraisal tools to provide study design-specific guidance on validity appraisal. Content analysis showed similar methodological and generalizability concerns were raised by both groups; however, the MetaQAT appraisers commented more extensively on applicability to public health practice.

Conclusions: Critical appraisal tools designed for clinical medicine have limitations for use in the context of public health. The meta-tool structure of the MetaQAT allows for rigorous

* Corresponding author. Dalla Lana School of Public Health, University of Toronto, Health Sciences Bldg, 6th floor, 155 College St., Toronto, Ontario M5T 3M7, Canada. Tel.: +1 416 978 6064.

E-mail addresses: laura.rosella@utoronto.ca (L. Rosella), Carolyn.bowman@oahpp.ca (C. Bowman), Beata.pach@oahpp.ca (B. Pach), Sarah.morgan@oahpp.ca (S. Morgan), Tiffanyfitzpatrick@ossu.ca (T. Fitzpatrick), Vivek.goel@utoronto.ca (V. Goel).
<http://dx.doi.org/10.1016/j.puhe.2015.10.027>

0033-3506/© 2015 The Authors. Published by Elsevier Ltd on behalf of The Royal Society for Public Health. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

appraisal, while allowing users to simultaneously appraise the multitude of study designs relevant to public health research and assess non-standard domains, such as applicability.

© 2015 The Authors. Published by Elsevier Ltd on behalf of The Royal Society for Public Health. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Introduction

A standard approach to critical appraisal is essential to ensure rigour and transparency. At a general level, critical appraisal is the process of evaluating the trustworthiness, value, and relevance of material¹ and is often aided by a tool, intended to guide the depth and breadth while increasing consistency and transparency.^{1–6} Tools employ a variety of techniques including question and answer format, checklist format, numeric scales and summary scores. Additionally, companion tools may be used in the company of other design- or method-specific tools to assess specific methodological elements. Overwhelmingly, appraisal tools have been developed in the context of evidence-based medicine (EBM); wherein, quality is predominantly determined according to the traditional study-design hierarchy and assessments of ‘risk of bias’.^{2–4} To make evidence-informed decisions in public health, evidence must be approached differently. Notably, the traditional evidence hierarchy becomes less relevant as randomised controlled trials, the most highly ranked form of evidence in EBM, are often not feasible or ethical. Therefore, the merit of specific study designs should be considered within the context of the question. This approach more appropriately addresses the heterogeneity of both evidence and enquiries in public health.^{2,5,6} Further, a different scope of considerations is required for appraisal in public health.² Critical appraisal is traditionally focused on ‘risk of bias’, which is highly appropriate for its origins in EBM, where high internal validity is important to prove the efficacy of an intervention. In public health, however, differences in setting or context are also vital considerations. Therefore, the critical appraisal approach must be able to adapt accordingly, given study designs and applications relevant to public health research and practice.

A useful appraisal tool requires a number of features. Foremost, it must be able to adapt to a variety of purposes: systematic review, topic enquiries, report updates, single study appraisal, etc. The tool should be appropriate for appraising different types of public health evidence: experimental, observational, reviews, qualitative, grey literature, etc. Third, it must apply across the spectrum of public health topics, regardless of maturity; e.g. developing health issues or emerging infectious diseases. Additionally, it must be suitable for assessing either one item in isolation or reviewing multiple studies. Finally, it must create a detailed record of the process, so that any inherent subjectivity of appraising is made transparent and may be evaluated and communicated. There are several challenges in developing a quality appraisal tool to meet the above criteria. Most tools are design-specific and focus only on ‘risk of bias’.^{7–11} The model of a ‘one tool fits all’ approach, however, is not only unlikely but also does not

make the best use of the existing science of quality appraisal, nor does it adapt in the way articulated in the aforementioned goals. However, this limitation may be overcome through the use of a meta-tool; a tool which orients the user to the appropriate use of several appraisal tools and places them within a larger framework to guide their use (Box 1).

Thus, the meta-tool allows for tools with disparate approaches to be used within the same process, providing consistency and transparency to the appraisal process. Further, this approach allows the appraisal framework to be expanded and relevant to public health.¹² The meta-tool allows for wide applicability and a high degree of flexibility, while maintaining a rigorous and transparent process that is required to conduct a detailed assessment.

The aim of this project was to develop a critical appraisal meta-tool that would meet the contemporary needs of public health researchers and practitioners who appraise and synthesise evidence for the purpose of recommendations, development of public health action, or scientific research.

Methods

Search strategy

Two information professionals (SM, BP) undertook a literature search with the goal of finding critical appraisal tools commonly used in public health. An initial MEDLINE search returned one relevant systematic review.⁷ The review reference list and subsequently those studies’ reference lists were hand-searched to identify additional tools, reviews, and other relevant documents. A second Medline search was conducted to identify tool comparisons, evaluations, and reviews published subsequent to the initial systematic review.

A grey literature search was conducted of the following organizational websites: The National Collaborating Centre for Methods and Tools (NCCMT) repository of critical appraisal tools, the Ontario Public Health Libraries Association (OPHLA), the Canadian Agency for Drugs and Technology in Health (CADTH), the Joanna Briggs Institute, the Cochrane Collaboration, the National Institute for Health and Clinical Excellence (NICE) and the European Centre for Disease Prevention and Control (ECDC). Library resources of

Box 1.

Definition of meta-tool for quality appraisal.

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.

well-known North American and Australian universities that grant degrees in medicine and public health were also searched: University of Toronto, McMaster University, Memorial University, Yale University, John Hopkins University, and the University of South Australia.

Selection criteria

The definition of tool was applied broadly; schemas and critical appraisal process documents were included. Tools were excluded if they did not address internal validity. Tools were included if they addressed a study design relevant to public health practice.

Comparison of tool items and development of quality framework

The selected critical appraisal tools were compared according to individual items; similarities were collated. The OPHLA critical appraisal guide,¹³ which builds on content created by the Public Health Research, Education, and Development Program (PHRED),¹⁴ was used as a starting point to categorize concepts of quality for public health evidence.

Once all of the components of the tool were mapped, the remaining questions on the tools were examined to determine whether they were relevant to appraising evidence in a public health context. The overall framework was developed from the domains that emerged from this mapping process. Preference was given to tools with documented validity and reliability. To determine the level of acceptability and adoption, the authors considered the presence of the tool on university library recommended lists and evidence of use in public health research.

Feedback sessions

The draft tool was tested during three feedback sessions. Participants were members of a large public health organization, who regularly provide technical and scientific advice on a wide range of public health issues, and included various levels of training and subject matter expertise.

One article was selected by the authors for the group to appraise¹⁵ and participants were invited to bring a second article from their area of work. Participants appraised the articles using the draft tool, followed by a group discussion focused on the interpretation of the tool elements. Three authors (LR, BP, SM) recorded feedback; relevant suggestions were identified and revisions were made accordingly.

Pilot testing and evaluation

The revised Meta Quality Assessment Tool (MetaQAT) was then piloted within several scientific teams beginning in November 2013. Pilot users were invited to share feedback in interviews and data was analyzed thematically.

Validation

Several aspects of validity were considered: face validity (appears to measure the concept it reports to measure), content

validity (covers all aspects of the concept it reports to measure), and criterion validity (ability to classify according to an agreed standard measure of the concept).¹⁶ Face validity was assessed by consulting with senior scientists experienced in critical appraisal. Content validity was established during the development process, when the content of relevant tools was compared and mapped according to a standard source. This process ensures that the MetaQAT framework covers all aspects of critical appraisal addressed by existing tools. In addition, the framework was compared with the Heller framework for appraisal in public health.² Since no gold standard tool exists,⁷ criterion validity was assessed by expert assessment of study quality.

A systematic process was designed to test validity. Specifically, three journal articles that reflect a range of public health issues were selected (Table 2).^{17–19} In order to remove variation in journal characteristics such as quality and reporting standards, all articles were selected from the same journal and year. Academic research staff, experienced in critical appraisal, were divided into two groups and asked to examine the articles for quality and report on the strengths and weaknesses; one used a generic critical appraisal feedback form and the other used MetaQAT. The generic form consisted of the following appraisal prompts: ‘What are the strengths of this study?’, ‘What are the weaknesses of this study?’, ‘Comment on sources of bias’, and ‘Comment on overall quality’.

The unstructured critical appraisal forms were collected and the content coded by CB, with TF acting as an independent second coder.²⁰ The codes were grouped by article and were compared across reviewers. Reviewer responses were discussed (CB, TF, LR) in order to decipher broad meanings until consensus was reached. The resulting codebook was reviewed by the study team, with only minor changes to ensure consistent coding of epidemiological concepts. The MetaQAT set of responses were then coded using the codebook. Code groupings evolved and new codes were added to match concepts expressed by group assigned the MetaQAT tool. The coding was summarized both by participant and appraisal group.

Content analysis was conducted using MAXQDA version 11.

Results

Search and selection of tools

Thirty-six relevant tools were identified and included in the item mapping process^{2,8,9,13,21–52} and included tools for systematic reviews, randomized controlled trials, observational studies, cohort studies, economic evaluations, needs analyses, clinical guidelines, user studies, quantitative studies (generic), qualitative studies, and mixed methods studies.

Item mapping outcome – critical appraisal framework

Through the mapping, a four-domain appraisal framework was developed, consisting of relevancy, reliability, validity, and applicability (Fig. A1). The tool provides generic guidance on how to assess validity; instead, the user is guided to the

Table 1 – Design-specific companion tools.

Research design	Recommended critical appraisal tool
Systematic reviews and meta-analyses	Assessing the Methodological Quality of Methodological Reviews (AMSTAR) ⁹
Cohort studies	Critical Appraisal Skills Program (CASP): Cohort Studies Checklist ²⁶
Case control studies	Critical Appraisal Skills Program (CASP): Case Control Studies Checklist ²⁵
Economic evaluation studies	Critical Appraisal Skills Program (CASP): Economic Evaluation Studies Checklist ²⁷
Non-randomised controlled trials	Transparent Reporting of Evaluations with Non-randomized Designs (TREND) ⁸
Randomised controlled trials	Consolidated Standards of Reporting Trials CONSORT Statement ⁴⁷
Mixed methods research	Evaluation Tool for Mixed Methods Studies ⁴²
Qualitative research	McMaster Critical Review Form – Qualitative Studies ⁴⁰ Guidelines for Critical Review Form: Qualitative Studies
Clinical guidelines	Appraisal of Guidelines, Research and Evaluation (AGREE II) Instrument ²²

companion set of existing, design-specific tools to provide detailed questions to augment the generic tool (Table 1). It is intended that the list of companion tools will continue to evolve as new needs are identified and new tools are developed. The fourth domain, applicability, is vital to the utility of the tool in public health. This domain asks the user to consider the different ways that the evidence could be used in public health; asking key questions to ensure all relevant public health concepts, such as equity, have been considered. This encourages the user to think beyond direct generalization of the results and consider the application of the evidence more broadly. Each domain consists of a main question with one or two parts, including several supplementary ‘hint’ questions meant to guide the user. Space is provided to record the answer to each main question. The answers are written in long form, with optional tick boxes corresponding to yes, no, unclear, or N/A. The long-form answer provides space for documentation of the appraisal process; allowing response reasoning to be recorded so that they can be later referred to and evaluated. Relatedly, because the tool is intended to be flexible and adaptable it is not prescriptive; many of the questions in the main tool are intentionally broad and open ended.

Feedback sessions

Following feedback from the first two sessions, revisions were made to increase usability and clarity. Participants of the third feedback session had fewer questions and suggestions regarding the revised tool; indicating that the revised version was more acceptable to participants. Following the third session, the authors (LR, BP, SM) discussed the three sessions and made any final revisions.

Evaluation

Preliminary results were used to clarify definitions of the domains and reorganize questions to better align with the domains. The main change concerned the terminology used, which was revised to match accepted research usage of the terms validity and reliability. The guide was rewritten to include explanations and instructions for each part of the tool.

Validation

A total of six summary tables (Appendix A1) were prepared. There was a high degree of agreement between the two groups on issues pertaining to validity; both groups of appraisers identified and commented on issues of design, measurement, confounding, selection bias, and external validity (Table 3). Comments notably differed within the other domains, particularly applicability. Furthermore, the MetaQAT group tended to highlight specific items that were not reported or not clear, potentially affecting both the overall appraisal and bias assessment between MetaQAT and generic form users. These differences are likely a result of the explicitness of the relevancy and applicability domains within MetaQAT.

Discussion

This study describes the rationale, development and validation of a new quality appraisal meta-tool for public health. MetaQAT expands the process of appraisal from being primarily focused on internal validity, to consider both the internal and external validity within the wider context of application to public health practice. Importantly, we present a transparent development and validation process, which is notably missing from most quality appraisal documentation and is an important aspect of deciding on the appropriateness of a given tool. When searching for an appropriate tool for use in our organization, we found ourselves facing two opposing needs. First, the tool must be flexible enough to accommodate the variety of enquires relevant to public health. This strongly suggested a generic tool would be best. Alternatively, the need for a rigorous process dictated the necessity of design-specific instruments. The meta-tool concept came about as a way to marry these two disparate needs. This conflict has been noted by others,^{2,12} and the potential for companion tools to bridge this gap has been previously suggested.¹² This approach was successful in creating a tool that is both flexible and specific, enabling the user to appraise all types of evidence using a single tool, without sacrificing the level of detail necessary for a truly rigorous appraisal.^{2,12} Therefore, it was decided to use a generic tool, accompanied by a set of companion tools for in-depth validity assessment. The mapping process ensured that

Table 2 – Summary of studies used in validation study.

Paper reference	Objective	Setting/Data	Study summary
Selassie et al., <i>AJE</i> 23 (2013) 750–756 ¹⁹	To study risk factors for severity and incidence of traumatic brain injury (TBI)	Surveillance data on injury from hospital in the United States	<ul style="list-style-type: none"> • Incidence of TBI from hospital data was reported and tracked over time • Authors noted an increase in sport-related TBI and high association with sport and off-road vehicular sport • The increase and high association with sport suggest need for increased surveillance and prevention efforts • Both adult proxy recall and household immunisation records have reasonable accuracy for classifying HPV vaccination status
Ojha et al., <i>AJE</i> 23 (2013) 281–285 ¹⁸	To assess the accuracy of Human Papilloma Virus (HPV) vaccination status based on adult proxy recall and household immunisation records	Nationally representative survey data on immunisation and corresponding provider data in Canada	<ul style="list-style-type: none"> • Small association was found between both prenatal and childhood infectious and prevalence of obesity later in life • Unmeasured confounding may explain the associations
Cocoros et al., <i>AJE</i> 23 (2013) 307–313 ¹⁷	To examine the relationship between early life infectious and adult obesity	Cohort study of males who underwent mandatory army fitness evaluations in Denmark	

the framework was consistent with the coverage of quality assessment in existing appraisal tools. Further development of the framework involved clarifying the meaning and coverage of each section through user feedback and interviews.

MetaQAT is uniquely structured and, therefore, less comparable to other tools. To the best of our knowledge, there is only one other appraisal tool designed specifically for use in public health,² although it is designed for individual studies and was not validated, but piloted. Similar to MetaQAT, this checklist has included additional aspects relevant to appraising public health evidence, namely ‘transferability’. Additionally, it is meant to be used within an alternative hierarchy of evidence specific to public health, and accompanied by the use of companion tools. Unlike MetaQAT, however, this checklist does not explicitly guide the user in assessing applicability; nor does it provide the user with a set of companion study design-specific tools. Further, this checklist was developed to be used in the appraisal of single items, whereas MetaQAT may be used to appraise single or multiple sources of evidence. Moreover, despite being deemed ‘useful’ by respondents during pilot examination, the checklist did not demonstrate improved performance over other tools.² Validation of MetaQAT suggests notable enhancements, particularly regarding references to applicability in practice.

Validation of critical appraisal tools is difficult as there is no established gold standard for any study design.⁷ We decided to use a qualitative approach to validation. While it is an innovative approach, we felt it best suited our situation. Using content analysis allowed us to understand and categorise the content of the appraisals and make meaningful comparisons between the two participant groups. Unlike most appraisal tools, MetaQAT does not include a numeric score and any attempt at quantifying appraisals done using MetaQAT is discouraged. Through our literature review and the tool's development, we recognized that the practice of assigning numeric scores is particularly unsuited to appraising public health evidence. Numeric scores obscure key information regarding the strengths and weaknesses of the evidence for public health practice. Even in situations where a summary score would have some degree of utility, their use should be discouraged as numeric summary scores have been shown to be unreliable.^{2,12,53–56} Further, the question and answer format of MetaQAT is designed to make it accessible, particularly for novice appraisers. Including an optional tick box to augment the long-form answer is meant to help orient new users who are more familiar with that format. However, the long-form answers will contain the key strengths and weaknesses of the evidence, and therefore, the information needed for thoughtful application of the evidence. Quality assessments should be made using the framework as a whole.

There are some limitations to be aware of when considering this study. Firstly, the search was initiated from an existing systematic review of appraisal tools. This was done because it was an efficient and practical way to begin identifying potential tools. We believe that the subsequent Medline, grey literature search and snowball searching from the original review resulted in an acceptable coverage of existing

Table 3 – Highlights of similarities and differences between the MetaQAT (M) and generic form (F) appraisal group assessments.

Study	Similarities in main issues	Differences in main issues
Selassie et al.	<ul style="list-style-type: none"> • Multiple sources of selection bias: <ul style="list-style-type: none"> → Outcome measure not specific to sport, resulting in misclassification: ‘e-codes that mention sport are limited so that only 1/3 of sports related injuries are identifiable’ M7 → Large exclusions based on activity (‘excluded bike-related TBI’ M3) and treatment location (‘missed TBI from private physician offices, urgent care facilities’ M7, ‘issues of access to care for poorer individuals’ M1) • Reliability and accuracy of hospitalisation data and use of validated measure: <ul style="list-style-type: none"> → ‘Hospital-based data ... and national administrative sources of population data should be considered reliable’ M1 → ‘Objective validated outcome data (administrative data)’ F1 	<ul style="list-style-type: none"> • MetaQAT group reported differences related to study setting and context, such as types of sports and sport culture, which would affect the findings <ul style="list-style-type: none"> → ‘Would people in Ontario be just as likely to play football like the study population in South Carolina? It may be hockey in Ontario.’ M7 • MetaQAT group considered application within public health context as a whole <ul style="list-style-type: none"> → ‘Severe TBI affect only small fraction of population. Focus on TBI in general/ any sports TBI may be more relevant to public health’ M2 • Ethics procedures noted by MetaQAT group only
Ojha et al.	<ul style="list-style-type: none"> • Significant selection bias: <ul style="list-style-type: none"> → Multiple exclusions, not clearly documented: ‘lack of sufficient descriptions on the percentage of subjects remained in the study after applying a number of exclusion criteria’ F4 → ‘Important group excluded from the study—those who do not have a regular health care provider’ M4 • Differences between analysis groups <ul style="list-style-type: none"> → ‘Noted demographic differences between groups’ F3 and potential for further unmeasured differences • Important public health topic for study <ul style="list-style-type: none"> → ‘Accuracy of these proxy measure is important to program planning’ M1, ‘applicable to a wide range of researchers and practitioners’ F2 • Appropriate study design using nationally representative sample, applied population weights • Appropriate use of gold standard measure 	<ul style="list-style-type: none"> • The MetaQAT group commented more extensively and more specifically on application, e.g. comparison of population to local population, note to consider the impact of using one dose of vaccine vs full coverage <ul style="list-style-type: none"> → ‘Study population likely similar enough to Ontario population’ M6 • Ethics procedures noted by MetaQAT group only
Cocoros et al.	<ul style="list-style-type: none"> • Residual and uncontrolled confounding significant sources of bias <ul style="list-style-type: none"> → ‘Many opportunities for residual confounding’ F1 → ‘Uncontrolled confounders draw resulting associations into question’ M1 • External validity and selection bias issues with very specific population <ul style="list-style-type: none"> → ‘Study results may not be generalisable to women or outside this narrow age range, or people with health problems that would have exempted them from the conscription examination’ F2 • Vague exposure measurement <ul style="list-style-type: none"> → ‘Limited scope, acute respiratory infections are very common in US, most not treated clinically’ M2 	<ul style="list-style-type: none"> • MetaQAT group were more likely to comment on the applicability of the evidence and relate to public health setting <ul style="list-style-type: none"> → ‘Not all that useful in the public health context. Difficult to do a lot more about reasons for maternal and child hospitalisations. If this was done, effect on obesity would be small’ M6 • Ethics procedures noted by MetaQAT group only

Table 3 – (continued)

Study	Similarities in main issues	Differences in main issues
	<ul style="list-style-type: none"> → 'Exposure (hospitalisation for infection) is a poor proxy for what they are truly interested in (infection)' F1 • Small effect size → 'Effect is so small' M6 → 'Very wide confidence intervals for any small effect size and therefore generally unconvincing results' F3 	

critical appraisal tools. A recent review of appraisal tools used a similar strategy as a practical and time-saving approach.⁵⁷ Additionally, another large review of critical appraisal tools found that an untraditional approach was required to identify appraisal tools.¹⁰ Secondly, the primary use of MetaQAT is to guide users in summarising the strengths and weaknesses of individual items of evidence. By allowing the same tool to be used across all types of evidence, MetaQAT overcomes the challenges associated with summarising the collective appraisal of a body of evidence. However, the challenge of assessing a body of evidence goes further than the question of which tool is used, and further work is required to fully address this need. Finally, we recognise the need for further validation and plan to conduct work outside of our organisation as the next step in establishing the validity of MetaQAT.

Conclusions

The MetaQAT fills a critical gap in appraisal tools available for use in public health research and practice. This paper provides a clear description of the development, evaluation and validation of this new tool. Validation among a group of public health researchers and practitioners suggests that appraisals of evidence completed using MetaQAT are comparable with judgements of quality made using generic appraisal guides. Appraisals completed using MetaQAT, however, expand quality assessment to include considerations relevant to public health, such as the application of the evidence to local context. It is our hope that MetaQAT will support the use of evidence beyond our organization, filling the critical appraisal tool gap currently experienced by the wider public health community.

Author statements

Ethical approval

Ethics for this study was granted by the Public Health Ontario Ethics Research Board (2013-047.01).

Funding

None declared.

Competing interests

None declared.

REFERENCES

- Burls A. *What is critical appraisal?* [Internet]. 2nd ed. Newmarket, UK: Hayward Group [cited 2015 Feb 10]. Available from: http://www.medicine.ox.ac.uk/bandolier/painres/download/whatis/what_is_critical_appraisal.pdf; 2009.
- Heller RF, Verma A, Gemmel I, Harrison R, Hart J, Edwards R. *Critical appraisal for public health: a new checklist.* *Public Health* 2008;122(1):92–8.
- Brownson RC, Fielding JE, Maylahn CM. *Evidence-based public health: a fundamental concept for public health practice.* *Annu Rev Public Health* 2009;30:175–201.
- SUNY Downstate Medical Center, Medical Research Library of Brooklyn. *Guide to research methods: the evidence pyramid* [Internet]. Brooklyn, NY: SUNY Downstate Medical Center [cited 2014 Sept 1]. Available from: <http://library.downstate.edu/EBM2/2100.htm>; 2004.
- Victoria CG, Habicht JP, Bryce J. *Evidence-based public health: moving beyond randomized trials.* *Am J Public Health*(3):400–5 [Internet]. [cited 2015 Feb 10]. Available from: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1448265/pdf/0940400.pdf>, 2004;94.
- Petticrew M, Roberts H. *Evidence, hierarchies, and typologies: horses for courses.* *J Epidemiol Community Health*(7):527–9 [Internet]. [cited 2015 Feb 10]. Available from: <http://jech.bmj.com/content/57/7/527.full.pdf+html>, 2003;57.
- Katrak P, Bialocerkowski AE, Massy-Westropp N, Kumar S, Grimmer KA. *A systematic review of the content of critical appraisal tools.* *BMC Med Res Methodol*:22 [Internet]. [cited 2014 Aug 7]. Available from: <http://www.biomedcentral.com/content/pdf/1471-2288-4-22.pdf>, 2004;4.
- Des Jarlais DC, Lyles C, Crepez NTREND Group. *Improving the reporting quality of nonrandomized evaluations of behavioral and public health interventions: the TREND statement.* *Am J Public Health*(3):361–6 [Internet]. [cited 2015 Feb 10]. Available from: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1448256/pdf/0940361.pdf>, 2004;94.
- Shea BJ, Grimshaw JM, Wells GA, Boers M, Andersson N, Hamel C, et al. *Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews.* *BMC Med Res Methodol*:10 [Internet]. [cited 2014 Aug 1]. Available from: <http://www.biomedcentral.com/content/pdf/1471-2288-7-10.pdf>, 2007;7.
- Bai A, Shukla VK, Bak G, Wells G. *Quality assessment tools project report.* Ottawa: Canadian Agency for Drugs and Technologies in Health; 2012.
- Sirriyeh R, Lawton R, Gardner P, Armitage G. *Reviewing studies with diverse designs: the development and evaluation of a new tool.* *J Eval Clin Pract* 2012;18(4):746–52.
- Voss PH, Rehfuss 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.

13. Ontario Public Health Libraries Association. *Critical appraisal of research evidence: a guide from the Ontario Public Health Libraries Association* [Internet]. Toronto, ON: Ontario Public Health Libraries Association [cited 2015 Feb 10]. Available from: <http://www.ophla.ca/resources.htm>; 2014.
14. Canadian Health Services Research Foundation. *Research use week (Northeastern Ontario): tools, strategies and stories of using evidence in rural and remote health services delivery and policy development* [Internet]. Ottawa, ON: Canadian Health Services Research Foundation [cited 2015 Feb 10]. Available from: http://www.cfhi-fcass.ca/migrated/pdf/event_reports/SudburyReport_E_FINAL.pdf; 2007.
15. Wada K, Oka-Ezoe K, Smith DR. Wearing face masks in public during the influenza season may reflect other positive hygiene practices in Japan. *BMC Public Health*:1065 [Internet]. [cited 2013 May 1]. Available from: <http://www.biomedcentral.com/content/pdf/1471-2458-12-1065.pdf>, 2012;12.
16. Singleton RA, Straits BC. *Approaches to social research*. New York: Oxford University Press; 2010:115–49. Chapter 5, Measurement.
17. Cocoros NM, Lash TL, Norgaard M, Farkas DK, DeMaria A, Sorensen HT. Hospitalized prenatal and childhood infections and obesity in Danish male conscripts. *Ann Epidemiol* 2013;23(6):307–13.
18. Ojha RP, Tota JE, Orrutt-Powell TN, Klosky JL, Ashokkumar R, Gurney JG. The accuracy of human papillomavirus vaccination status based on adult proxy recall or household immunization records for adolescent females in the United States: results from the National Immunization Survey-Teen. *Ann Epidemiol* 2013;23(5):281–5.
19. Selassie AW, Wilson DA, Pickelsimer EE, Voronca DC, Williams NR, Edwards JC. Incidence of sport-related traumatic brain injury and risk factors of severity: a population-based epidemiologic study. *Ann Epidemiol* 2013;23(12):750–6. Available from: <http://www.annepidemiol.com/content/pdf/1050-6288-23-12-750.pdf>.
20. Elo S, Kyngäs H. The qualitative content analysis process. *J Adv Nurs* 2008;62(1):107–15.
21. Booth A, Brice A. Column editors. *HILJ*. 2001;18(3):175–77.
22. Brouwers MC, Kho ME, Browman GP, Burgers JS, Cluzeau F, Feder G, et al. AGREE II: advancing guideline development, reporting and evaluation in health care. *CMAJ*(18):E839–42 [Internet]. [cited 2015 Feb 10]. Available from: <http://www.cmaj.ca/content/182/18/E839.full.pdf+html>, 2010;182.
23. Centre for Evidence-Based Medicine. *Critical appraisal worksheets: systematic review critical appraisal sheet* [Internet]. Oxford: Centre for Evidence-Based Medicine [cited 2015 Feb 10]. Available from: <http://www.cebm.net/critical-appraisal/>; 2014.
24. Centre for Evidence-Based Medicine. *Critical appraisal worksheets: therapy/RCT critical appraisal sheet* [Internet]. Oxford: Centre for Evidence-Based Medicine [cited 2015 Feb 10]. Available from: <http://www.cebm.net/critical-appraisal/>; 2014.
25. Critical Appraisal Skills Programme (CASP). *CASP: case control studies* [Internet]. Oxford: Critical Appraisal Skills Programme [cited 2015 Feb 10]. Available from: <http://www.casp-uk.net/>; 2010.
26. Critical Appraisal Skills Programme (CASP). *CASP: cohort study checklist* [Internet]. Oxford: CASP [cited 2015 Feb 10]. Available from: http://media.wix.com/ugd/dded87_36c5c76519f7bf14731ed1985e8e9798.pdf; 2013.
27. Critical Appraisal Skills Programme (CASP). *CASP: economic evaluations checklist* [Internet]. Oxford: CASP [cited 2015 Feb 10]. Available from: http://media.wix.com/ugd/dded87_ff37ab65474ac19e19ec31d5f2e4887d.pdf; 2013.
28. Critical Appraisal Skills Programme (CASP). *CASP: qualitative research checklist* [Internet]. Oxford: CASP [cited 2015 Feb 10]. Available from: http://media.wix.com/ugd/dded87_951541699e9edc71ce66c9bac4734c69.pdf; 2013.
29. Critical Appraisal Skills Programme (CASP). *CASP: randomized controlled trials checklist* [Internet]. Oxford: CASP [cited 2015 Feb 10]. Available from: http://media.wix.com/ugd/dded87_c7984187fdc74d25a3fba6c707383b15.pdf; 2013.
30. Critical Appraisal Skills Programme (CASP). *CASP: systematic reviews checklist* [Internet]. Oxford: CASP [cited 2015 Feb 10]. Available from: http://media.wix.com/ugd/dded87_37491d0241aa448a8f3d4ae17c869472.pdf; 2013.
31. Glynn L. EBLIP critical appraisal checklist [Internet]. Newfoundland: Memorial University of Newfoundland; n.d. [cited 2015 Feb 10]. Available from: <http://ebltoolkit.pbworks.com/f/EBLCriticalAppraisalChecklist.pdf>.
32. Grimmer K, Dizon JM, Milanese S, King E, Beaton K, Thorpe O, et al. Efficient clinical evaluation of guideline quality: development and testing of a new tool. *BMC Med Res Methodol*:63 [Internet]. [cited 2014 Dec 1]. Available from: <http://www.biomedcentral.com/content/pdf/1471-2288-14-63.pdf>, 2014;14.
33. Health Evidence. *Quality assessment tool – review articles* [Internet]. Hamilton, ON: Health Evidence [cited 2015 Feb 10]. Available from: http://www.healthevidence.org/documents/our-appraisal-tools/QA_tool&dictionary_18.Mar.2013.pdf; 2013.
34. Jackson R, Ameratunga S, Broad J, Connor J, Lethaby A, Robb G, et al. The GATE frame: critical appraisal with pictures. *Evid Based Med* 2006;11(2):35–8.
35. Joanna Briggs Institute. *JB I SUMARI: NOTARI* [Internet]. South Australia: Joanna Briggs Institute [cited 2015 Feb 10]. Available from: <http://joannabriggs.org/sumari.html>; 2014.
36. Koufogiannakis D, Booth A, Brett A. *ReLIANT: reader's guide to the literature on interventions addressing the need for education and training*. *Libr Inf Res* 2006;30(94).
37. Centre for Evidence-Based Medicine. *Harm critical appraisal worksheet* [Internet]. Ottawa, ON: KT Clearinghouse [cited 2015 Feb 10]. Available from: <http://ktclearinghouse.ca/cebm/teaching/worksheets/harm/>; 2014.
38. Centre for Evidence-Based Medicine. *Systematic review (of therapy) critical appraisal worksheet* [Internet]. Ottawa, ON: KT Clearinghouse [cited 2015 Feb 10]. Available from: <http://ktclearinghouse.ca/cebm/teaching/worksheets/sr/>; 2014.
39. Law M, Stewart D, Pollock N, Letts L, Bosch J, Westmorland M. *Critical review form – quantitative studies* [Internet]. Hamilton, ON: McMaster University [cited 2015 Feb 10]. Available from: <http://www.srs-mcmaster.ca/Portals/20/pdf/ebp/quantreview.pdf>; 1998.
40. Letts L, Wilkins S, Law M, Stewart D, Bosch J, Westmorland M. *Critical review form – qualitative studies. Version 2.0* [Internet]. Hamilton, ON: McMaster University [cited 2015 Feb 10]. Available from: http://www.srs-mcmaster.ca/Portals/20/pdf/ebp/qualreview_version2.0.pdf; 2007.
41. Long AF, Godfrey M, Randall T, Brett A, Grant MJ. *HICPRDU evaluation tool for quantitative studies* [Internet]. Leeds, UK: University of Leeds [cited 2015 Feb 10]. Available from: <http://usir.salford.ac.uk/12969/>; 2002.
42. Long AF, Godfrey M, Randall T, Brett A, Grant MJ. *HICPRDU evaluation tool for mixed methods studies* [Internet]. Leeds, UK: University of Leeds [cited 2015 Feb 10]. Available from: <http://usir.salford.ac.uk/13070/>; 2002.
43. MacAuley D. *READER: an acronym to aid critical reading by general practitioners*. *Br J Gen Pract*(379):83–5 [Internet]. [cited 2015 Feb 10]. Available from: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1238789/pdf/brjgenprac00035-0039.pdf>, 1994;44.
44. Maher CG, Sherrington C, Herbert RD, Moseley AM, Elkins M. *Reliability of the PEDro scale for rating quality of randomized controlled trials*. *Phys Ther* 2003;83(8):713–21.

45. Moher D, Liberati A, Tetzlaff J, Altman DGPRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *BMJ*:b2535 [Internet]. [cited 2015 Feb 10]. Available from: <http://www.bmj.com/content/339/bmj.b2535.long>, 2009;339.
46. Pluye P, Gagnon MP, Griffiths F, Johnson-Lafleur J. A scoring system for appraising mixed methods research, and concomitantly appraising qualitative, quantitative and mixed methods primary studies in mixed studies reviews. *Int J Nurs Stud* 2009;46(4):529–46.
47. Schulz KF, Altman DG, Moher DCONSORT Group. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. *PLoS Med*(3):e1000251 [Internet]. [cited 2015 Feb 10]. Available from: <http://www.plosmedicine.org/article/fetchObject.action?uri=info:doi/10.1371/journal.pmed.1000251&representation=PDF>, 2010;7.
48. Scottish Intercollegiate Guidelines Network. *Methodology checklist 2: randomized controlled trials* [Internet]. Edinburgh: Scottish Intercollegiate Guidelines Network [cited 2015 Feb 10]. Available from: <http://www.sign.ac.uk/methodology/checklists.html>; 2014.
49. Scottish Intercollegiate Guidelines Network. *Methodology checklist 1: systematic reviews and meta-analyses* [Internet]. Edinburgh: Scottish Intercollegiate Guidelines Network [cited 2015 Feb 10]. Available from: <http://www.sign.ac.uk/methodology/checklists.html>; 2014.
50. Scottish Intercollegiate Guidelines Network. *Methodology checklist 3: cohort studies* [Internet]. Edinburgh: Scottish Intercollegiate Guidelines Network [cited 2015 Feb 10]. Available from: <http://www.sign.ac.uk/methodology/checklists.html>; 2014.
51. Scottish Intercollegiate Guidelines Network. *Methodology checklist 4: case-control studies* [Internet]. Edinburgh: Scottish Intercollegiate Guidelines Network [cited 2015 Feb 10]. Available from: <http://www.sign.ac.uk/methodology/checklists.html>; 2014.
52. AGREE Collaboration. Development and validation of an international appraisal instrument for assessing the quality of clinical practice guidelines: the AGREE project. *Qual Saf Health Care*(1):18–23 [Internet]. [cited 2015 Feb 10]. Available from: <http://qualitysafety.bmj.com/content/12/1/18.full.pdf+html>, 2003;12.
53. 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.
54. Higgins JPT, Green S, editors. *Cochrane handbook for systematic reviews of interventions* [Internet]. Version 5.1.0 [updated March 2011]. Cochrane Collaboration; 2011 [cited 2015 Feb 10]. Section 8.3.3, Quality scales and Cochrane reviews. Available from: <http://handbook.cochrane.org>.
55. Juni P, Witschi A, Bloch R, Egger M. The hazards of scoring the quality of clinical trials for meta-analysis. *JAMA*(11):1054–60 [Internet]. [cited 2015 Feb 10]. Available from: <http://jamanetwork.com/data/Journals/JAMA/4692/JOC81641.pdf>, 1999;282.
56. Wendt O, Miller B. Quality appraisal of single-subject experimental designs: an overview and comparison of different appraisal tools. *Educ Treat Child* 2012;35(2):235–68.
57. Harder T, Takla A, Rehfuess E, Sanchez-Vivar A, Mtysiak-Klose D, Echmanns 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*:69 [Internet]. [cited 2014 Dec 1]. Available from: <http://www.biomedcentral.com/content/pdf/1471-2288-14-69.pdf>, 2014;14.

Appendix A. Supplementary data

Supplementary data related to this article can be found at <http://dx.doi.org/10.1016/j.puhe.2015.10.027>.