

Development of a framework to identify research gaps from systematic reviews

Karen A. Robinson*, Ian J. Saldanha, Naomi A. Mckoy

Division of General Internal Medicine, Department of Medicine, Johns Hopkins University, 1830 East Monument Street, Room 8069, Baltimore, MD, USA

Accepted 11 June 2011; Published online 19 September 2011

Abstract

Objective: Our objective was to develop a framework to identify research gaps from systematic reviews.

Study Design and Setting: We reviewed the practices of (1) evidence-based practice centers (EPCs), and (2) other organizations that conduct evidence syntheses. We developed and pilot tested a framework for identifying research gaps.

Results: Four (33%) EPCs and three (8%) other organizations reported using an explicit framework to determine research gaps. Variations of the PICO (population, intervention, comparison, outcomes) framework were most common. We developed a framework incorporating both the characterization of the gap using PICOS elements (also including setting) and the identification of the reason(s) why the gap exists as (1) insufficient or imprecise information, (2) biased information, (3) inconsistency or unknown consistency, and (4) not the right information. We mapped each of these reasons to concepts from three common evidence-grading systems.

Conclusion: Our framework determines from systematic reviews where the current evidence falls short and why or how the evidence falls short. This explicit identification of research gaps will allow systematic reviews to maximally inform the types of questions that need to be addressed and the types of studies needed to address the research gaps. © 2011 Elsevier Inc. All rights reserved.

Keywords: Research gaps; Systematic review; Evidence-based practice; Evidence-based research; Research priorities; Framework

1. Background and objective

Research should answer questions that matter. Setting a research agenda requires identifying questions that are important to patients, policy makers, and clinicians. A research agenda should also reflect the appropriate ways to answer the questions. The identification of where the evidence falls short and how the evidence falls short is essential to the development of important research questions and in providing guidance in how to address those questions.

The synthesis of existing evidence, through systematic reviews, has been widely recognized as the best way to inform decisions about health care. However, the full potential of systematic reviews has not been realized. The consideration of existing evidence often highlights important areas where deficiencies in information limit our ability to make decisions. When the ability of the systematic reviewer to draw conclusions is limited, we have called this a “research gap.” A research gap may be further developed, such as through stakeholder involvement in prioritization, into

“research needs.” Research needs are those areas where the gaps in the evidence limit decision making by patients, clinicians, and policy makers. The clear and explicit identification of research gaps is a necessary step in developing a research agenda, including decisions about funding and the design of informative studies. Yet, the identification of such gaps has not been completed in a systematic way. Instead, people seeking guidance from systematic reviews about useful future research are often frustrated by the lack of detail provided.

Clarke et al. [1] examined the 2,535 Cochrane reviews in Issue 4, 2005 of The Cochrane Library. Cochrane reviews include a separate section for review authors to discuss “Implications for Research.” However, Clarke et al. found that the characterization of the needs for future research was less than explicit. Twelve percent of the reviews failed to specify any of three basic components of a well-designed research question (population, intervention, outcome). Only about 17% included all three domains in describing gaps. Although not a purpose of the Clarke et al.’s study, it also appeared that reviews did not provide any description of how gaps were identified.

Scott et al. [2] surveyed 43 member organizations of the International Network of Agencies for Health Technology

* Corresponding author. Tel.: (001) 410-502-9216; fax: (001) 410-955-0825.

E-mail address: krubin@jhmi.edu (K.A. Robinson).

What is new?

Key finding:

A framework was developed to explicitly identify and characterize research gaps from systematic reviews.

What this adds to what was known?

Research gaps prevent systematic reviewers from making conclusions and, ultimately, limit our ability to make informed health care decisions. This framework characterizes where the evidence falls short and how or why the evidence falls short.

What is the implication and what should change now?

Use of the framework promotes an explicit and systematic method of identification and presentation of research gaps from systematic reviews.

Assessment (INAHTA) and classified two organizations of 12 responses as having a formal process for linking gaps from health technology assessments (HTAs) to the research funding process. In a review of the HTAs from their own organization, the authors noted that, although many of the HTAs reported that there was limited evidence, gaps were not specifically highlighted and discussion of the limitations of the evidence was embedded in the text [2,3].

We completed a search for articles describing methods for the identification of research gaps from systematic reviews (MEDLINE, April 22, 2010). We found five relevant studies [1,4–7]. Other than the Clarke et al. study, these studies were all conducted within a specific topic or disease area. In 2009, Chou et al. [4] evaluated the body of evidence for each research question related to the use of opioids for chronic noncancer pain, and considered research gaps as those questions with only “poor quality” evidence. Two of the articles used topic-specific organizing principles, including a care pathway [5] and a decision tree [7], to identify gaps as areas with limited or poor quality evidence.

We found limited information in the literature describing the development or use of frameworks for the identification of research gaps. Our objective was to develop a framework for the identification of research gaps from systematic reviews.

2. Methods

Detailed methods are provided in the report for this Evidence-based Practice Center project funded by the Agency for Healthcare Research and Quality (AHRQ) [8]. We first sought to identify any methods and frameworks being used by organizations to identify gaps from

systematic reviews. We conducted an audit of reports from Evidence-based Practice Centers (EPCs). Reports addressing clinical or health care services questions, published since 2008 by current EPCs, were considered. If an EPC had multiple reports that met these criteria, one report was randomly chosen for inclusion in our sample. We also contacted other organizations that conduct systematic reviews or related syntheses, such as technology assessments. These organizations were all current member organizations of INAHTA [9] or current members of the Guidelines International Network [10] from the United States, Canada, U.K., or Australia who conduct syntheses of evidence.

We developed a framework based on the results of the audit of current practices. We also developed a worksheet and a set of instructions to be used in the application of the framework. The draft framework was reviewed by two experts at our institution. We then pilot tested the revised framework using two randomly chosen systematic reviews from EPCs. Two team members independently applied the framework to each report. Process information, such as time for completion and any issues in the application of the framework, was collected. The number and type of gaps identified using our framework were compared with those presented in the future research sections of the reports.

3. Results

Twelve reports were included in our audit of EPC reports. All reports included some discussion of future research needs or gaps; in 75%, this was found in the discussion section of the report. None of the reports included a description of how research gaps or needs were identified. We determined that two of the 12 reports used an explicit framework to present the gaps. In both cases, the PICO format was used (population, intervention, comparison, outcomes).

We identified and contacted 64 organizations that conduct evidence syntheses. We received a response from 37 (58%). We classified four (11%) organizations as having a formal process for the identification of research gaps or needs; three of these used a framework. Two of the three organizations used PICO and one organization used the key questions from a guideline as a framework. Eleven other organizations reported using a formal process but nine of these did not meet our definition, and two used a formal process for identifying needs for evidence syntheses ($n = 1$) or guidelines ($n = 1$).

On the basis of our review and scan of current practices, we felt that a framework should comprise a description of where the evidence is inadequate and a description of the way(s) in which the evidence is inadequate. We developed a framework to be applied for each review question and to include two elements: (1) the characterization of the gaps and (2) the identification of the reason(s) for each research gap. The PICOS or PICO format is used by many organizations in developing questions, and was used to characterize

research gaps by projects we identified in the literature and by a number of the EPCs and other organizations. In applying the PICOS element of the framework, an investigator would provide details about the research gap in terms of population (P), intervention (I), comparison (C), outcome (O), and setting (S). Those elements that are inadequately addressed in the evidence base would be described. The other relevant elements will be apparent from the review question from which the research is derived. For research gaps that do not relate to a specific review question, all available elements of the research gap would be characterized.

The other concept of the framework is a classification of the most important reason(s) for the existence of the research gap. The reason(s) indicated would be those that most preclude conclusions from being made. Put another way, what would be needed to allow for conclusions to be made? The proposed classification of the reasons for research gaps includes:

- A. Insufficient or imprecise information,
- B. Biased information,
- C. Inconsistency or unknown consistency, and
- D. Not the right information.

To leverage work already being completed by review teams, we mapped each of these reasons for research gaps to concepts from three commonly used evidence grading systems: the United States Preventive Services Task Force (USPSTF) system [11], the Strength of Evidence (SOE) system used by EPCs [12], and the Grading of Recommendations Assessment, Development and Evaluation (GRADE) system [13].

The classification of reason(s) for gap are based on the following:

3.1. *Insufficient or imprecise information*

Insufficient information can arise if no studies are identified, if a limited number of studies are identified, or if the sample sizes in the available studies are too small to allow conclusions about the question of interest. If the information available in identified studies is insufficient to allow a conclusion or if the estimate of the effect (usually achieved from a meta-analysis) is imprecise there is a research gap. Precision is the degree of certainty surrounding the effect estimate. An imprecise estimate has been defined as one for which the confidence interval is wide enough to include both superiority and inferiority (i.e., the direction of effect is unknown), a circumstance that precludes a conclusion [12]. More broadly, imprecision could be considered to exist if the upper and lower levels of the confidence interval would result in different clinical decisions. Imprecision in the meta-analytic effect estimate may result as a consequence of a small number of studies in the meta-analysis or small sample sizes in included studies (leading to imprecision in individual study effect sizes). Where meta-analysis is not

conducted, precision of the individual studies should be evaluated. This category corresponds to the precision domain in the EPC SOE, GRADE and USPSTF both include consideration of this concept. USPSTF, for instance, considers the number and size of the studies that address a particular question, whereas the GRADE Working Group advise decreasing the grade of quality of evidence if data are “imprecise or sparse.”

3.2. *Biased information*

Various criteria exist for assessing the risk of bias of studies of different study designs. The aggregate risk of bias is contingent on the risk of bias of the individual studies [12]. In addition to considering methodological limitations of studies, the appropriateness of the study design should also be considered. The risk of bias of the body of evidence would also be considered here, such as the possibility of publication bias or other reporting biases (e.g., selective outcome reporting). Each of the grading systems incorporates elements of study design and the aggregate of the quality of the studies.

3.3. *Inconsistency or unknown consistency*

In the SOE system used by EPCs, consistency is defined as the degree to which reported effect sizes from included studies appear to go in the same direction [12]. The two elements are whether effect sizes have the same sign (same side of “no effect”) and whether the range of effect sizes is narrow. According to the GRADE system, consistency refers to the similarity of estimates of effect across studies, incorporating direction of effect, size of differences in effect, and the significance of the differences in effect size [13]. However, it should be kept in mind that a statistically significant effect size in one study and an effect size whose confidence interval overlaps null in another study do not necessarily constitute inconsistent results. Statistical measures of heterogeneity may be used to help in evaluating consistency. If there is only one available study, even if considered large sample size, the consistency of results is unknown [12].

3.4. *Not the right information*

There are a number of reasons why identified studies might not provide the right information for answering the question of interest. First, results from studies might not be applicable to the population and/or setting of interest. Secondly, the optimal or most important outcomes might not be assessed. For example, studies might only include surrogate or intermediate outcomes. Thirdly, the study duration might be too short and patients might not be followed-up for long enough duration to adequately assess some outcomes that might be most important. This reason for research gap maps to several different concepts in the grading systems. In the EPC SOE, applicability is included as an “other

pertinent issue” and directness is a required domain. This system also incorporates the consideration of surrogate vs. clinical outcomes. Directness is a key element, incorporating the elements of applicability and surrogate vs. clinical outcomes in GRADE. The USPSTF considers the question “To what extent are the results of the studies generalizable to the general U.S. primary care population and situation? (i.e., what is the external validity?).”

We designed a worksheet to facilitate the application of the framework for the identification and organization of research gaps during systematic reviews (see Box 1). Our aim was to design a simple, user-friendly worksheet to help investigators record research gaps. We envision that investigators would fill out this worksheet soon after the data synthesis phase, while in the process of writing the results section of the systematic review. Having just completed reviewing the evidence in detail, we believe that this is the ideal time for investigators to comprehensively and accurately identify individual research gaps.

After internal review, and resulting clarification of the framework, we pilot tested our framework on two randomly selected evidence reports not produced by our EPC [14,15]. It took an average of 3.5 hours per evidence report and we

identified a mean of 14.75 research gaps per evidence report. There were about two (14%) research gaps per evidence report, which could not be characterized using the PICOS component of the framework. These research gaps did not relate to a specific intervention or comparison, but instead related to prevalence, incidence, and the effect of certain factors on prevalence and incidence. These research gaps were thus abstracted in free-text form. There were some differences between what we abstracted and the future research sections of the report in terms of the level of detail. For instance, the future research sections in the EPC reports tended to make more general recommendations about research needs than we abstracted using the framework.

4. Conclusions

Systematic reviews are essential to the practice of “evidence-based research.” Health care research should begin and end with a systematic review [16–18]. A comprehensive and explicit consideration of the existing evidence is necessary for the development of an unanswered and answerable question, for the design of a study most likely to

Box 1 Research gaps abstraction worksheet (with examples)

Pilot Test of Framework – EPC Report I
Research Gap Worksheet

Completed by – IJS
Date – 08/31/2010
Page 1 of 8
Key Question Number – 1

Serial Number	Reason(s) for Gap* (Enter codes for ALL that apply)	POPULATION (P)	INTERVENTION (I)	COMPARISON (C)	OUTCOMES (O)	SETTING (S)	Free text of gap	Notes
1	C	Women with overactive bladder	Estradiol releasing vaginal ring	Estradiol vaginal pessary				
2	A	Women with overactive bladder	Sacral neuromodulation	Medical therapy				
3	C	Women with overactive bladder	Oxybutryin	Trospium				
4	C	Women with overactive bladder	Oxybutryin	Darifenacin				
5	C	Women with overactive bladder	Oxybutryin	Tolterodine				
6	A	Women with overactive bladder	Pelvic floor muscle training	Electric vaginal stimulation				

* Reasons for Gap -

- A. Insufficient (no studies/ limited number of studies/ small sample size(s)) or imprecise information
- B. Biased information (high risk of bias/ suboptimal study design)
- C. Inconsistent or unknown consistency results
- D. Not the right information (results not applicable/ optimal outcomes not assessed/ studies too short)

answer that question, and for the interpretation of the results of the study [19]. The identification of gaps from systematic reviews is one way to move toward evidence-based research.

We developed a framework to facilitate the identification and characterization of research gaps from systematic reviews. The framework provides for the classification of where the current evidence falls short and why or how the evidence falls short. Knowing where the gaps are and the reason(s) underlying their existence could help in the translation of these gaps into specific research needs, and subsequently, in the prioritization and design of the appropriate research to fill them. The proposed framework, and accompanying worksheet, may help in making systematic reviews a more reliable source of research gaps in two main ways. First, it facilitates a systematic process to identify and record research gaps during systematic reviews. This would also facilitate the discussion about research gaps between team members who might have written the results for different review questions. Second, use of the framework would enable investigative teams to write the future research section of an evidence report in a more organized and systematic manner. A proposed format for presenting research gaps that takes full use of the framework is provided in Box 2. Not surprisingly, given the use of PICOS structure, our proposed format for presentation is similar to that proposed by Brown et al. [20]. Further research on the presentation of research gaps could assess if other elements of the EPICOT+ template (Evidence, Population, Intervention, Comparison, Outcome, Time Stamp) might be useful for different stakeholders. Alternatively, research with stakeholders may determine a preference for more of a free-form, free-text method of presenting research gaps.

In identifying research gaps, we suggest that investigative teams decide a priori about the specificity of research gaps to be identified and presented. It is also important to decide a priori which reason will be selected when the

research gap arises because only one study is identified (i.e., insufficient information or unknown consistency). When identifying reasons why a research gap exists, it is helpful if team members pick the main reason(s) that prevented conclusions from being made and to be as specific as possible. This would provide the most useful guidance in designing the appropriate research to fill that gap.

Our framework calls for identifying the most important reason(s) for existence of research gaps. However, there may often be more than one main reason why a research gap exists. Team members could differ in their judgment of the relative importance of these reasons. Decisions on the relative importance of these reasons are often arbitrary. More research is needed to determine if a hierarchy or ranking system can be established to aid these decisions.

The application of the framework to identify research gaps by our investigative team was challenging. Much of this was because of our team being unfamiliar with the evidence reports and trying to retrospectively apply the framework. We suggest that the use of the framework will be more efficient if the same investigative team, which synthesizes the evidence applies the framework while writing the results. This can be empirically tested going forward. Also, if the body of evidence is graded, such as with USPSTF, GRADE, or the SOE system used by EPCs, teams can leverage work done in grading the evidence to identify research gaps.

Further evaluation is needed to see how the framework performs during the completion of a systematic review. A future project could have some members of the review team use the worksheet, and others not, to compare the process and outcome (i.e., future research section). In addition to further testing of implementation, the framework needs to be tested across a larger number and variety of reviews. This would enable determination of more formal metrics of instrument testing such as reliability. Given the use of the PICOS structure, it is not clear how the framework will

Box 2 Proposed format for presentation of research gaps

Key question number and key question topic

Research gap number

- Reason for gap
- Population (P)
- Intervention (I)
- Comparison (C)
- Outcomes (O)
- Setting (S)

Research question

Key question 1—What are the risks and benefits of oral diabetes agents (e.g., second-generation sulfonylureas and metformin) as compared with all types of insulin in women with gestational diabetes?

Research gap number 1

- Reason for gap—not the right information (long-term outcomes not assessed)
- Population (P)—women with gestational diabetes
- Intervention (I)—glipizide
- Comparison (C)—any insulin
- Outcomes (O)—long-term maternal outcomes (postpartum weight retention, obesity, development of type 2 diabetes), and long-term offspring outcomes (chronic diseases, e.g., obesity, type 2 diabetes)
- Settings (S)—any setting

Research question number 1

What are the effectiveness and safety of glipizide compared with any insulin in the treatment of gestational diabetes with regard to the long-term maternal outcomes (postpartum weight retention, obesity, development of type 2 diabetes) and long-term offspring outcomes (chronic diseases, e.g., obesity, type 2 diabetes)?

perform with nonintervention types of questions. Further refinement of the framework may be necessary to enable the use across different types of questions. This testing may also highlight other necessary modifications of the framework. For instance, our current work suggests that the category “not the right information” may be split into different concepts (i.e., applicability, appropriate outcomes, etc.) to ease application of the framework and promote a more clear presentation of the research gaps.

Further evaluation is also needed to determine if the gaps identified using the framework are different than those identified using current methods. This could be assessed by examining the number and type of gaps identified. Perhaps more meaningful would be the evaluation of the perceived usefulness of the gaps by potential stakeholders. It is not clear if the translation of research gaps to prioritized research needs, in which issues such as importance and feasibility are considered, would be a more efficient or comprehensive process if our framework was used and presented as part of systematic reviews.

In synthesizing the available evidence, systematic reviews inform health care decisions for patients, policy makers, and clinicians. Systematic reviews can and should also be an invaluable source for the identification of research gaps, thus informing the development of research agendas. This potential impact of systematic reviews has not been fully realized. Our framework provides for the systematic identification and characterization of research gaps from systematic reviews. This explicit identification of research gaps will allow systematic reviews to maximally inform the types of questions that need to be addressed and the type of studies needed to address the research gaps.

Acknowledgments

The authors are grateful to Jodi Segal, MD, MPH, and Steve Goodman, MD, PhD, of the Johns Hopkins University School of Medicine for their expert review of the framework, the worksheet, and the accompanying instructions. They thank the peer reviewers for the report of the AHRQ-funded Evidence-based Practice Center project. They also thank Laura Barnes, BS, for her assistance in abstracting data from EPC evidence reports.

This project was funded under Contract No. HHSA 290-2007-10061-I from the Agency for Healthcare Research and Quality, U.S. Department of Health and Human Services. The authors of this report are responsible for its content. Statements in the report should not be construed as endorsement by the Agency for Healthcare Research and Quality or the U.S. Department of Health and Human Services.

References

- [1] Clarke L, Clarke M, Clarke T. How useful are Cochrane reviews in identifying research needs? *J Health Serv Res Policy* 2007;12:101–3.
- [2] Scott A, Moga C, Harstall C. Using HTA to identify research gaps: a pilot study. 2006; [Web Page]. Available at http://www.ihe.ca/documents/HTA_Initiative24.pdf. Accessed May 17, 2011.
- [3] Scott NA, Moga C, Harstall C, Magnan J. Using health technology assessment to identify research gaps: an unexploited resource for increasing the value of clinical research. *Health Policy* 2008;3(3):e109–27.
- [4] Chou R, Ballantyne JC, Fanciullo GJ, Fine PG, Miaskowski C. Research gaps on use of opioids for chronic noncancer pain: findings from a review of the evidence for an American Pain Society and American Academy of Pain Medicine clinical practice guideline. *J Pain* 2009;10:147–59.
- [5] Shepherd J, Briggs J, Payne L, Packer C, Kerridge L, Ashton-Key M. Setting the future policy agenda for health technology assessment: a specialty mapping approach. *Int J Technol Assess Health Care* 2007;23:405–13.
- [6] Johnson NP, Proctor M, Farquhar CM. Gaps in the evidence for fertility treatment—an analysis of the Cochrane Menstrual Disorders and Subfertility Group database. *Hum Reprod* 2003;18:947–54.
- [7] de Vet HC, Kroese ME, Scholten RJ, Bouter LM. A method for research programming in the field of evidence-based medicine. *Int J Technol Assess Health Care* 2001;17:433–41.
- [8] Robinson KA, Saldanha IJ, McKoy NA. Frameworks for determining research gaps during systematic reviews. No. xx. (Prepared by the Johns Hopkins University Evidence-based Practice Center under Contract No. HHSA 290-2007-10061-I.). Rockville, MD: Agency for Healthcare Research and Quality; <Month Year>. Available at www.effectivehealthcare.ahrq.gov/reports/final.cfm. Accessed May 17, 2011.
- [9] International Network of Agencies for Health Technology Assessment (INAHTA). List of members. [Web Page]. Available at <http://www.inahta.org/Members/Contact-database/>. Accessed May 17, 2011.
- [10] Guidelines International Network (G-I-N). Members around the world. [Web Page]. Available at <http://www.g-i-n.net/membership/members-around-the-world>. Accessed May 17, 2011.
- [11] Sawaya GF, Guirgis-Blake J, LeFevre M, Harris R, Petitti D. Update on the methods of the U.S. Preventive Services Task Force: estimating certainty and magnitude of net benefit. *Ann Intern Med* 2007;147:871–5.
- [12] Owens DK, Lohr KN, Atkins D, Treadwell JR, Reston JT, Bass EB, et al. AHRQ series paper 5: grading the strength of a body of evidence when comparing medical interventions—agency for healthcare research and quality and the effective health-care program. *J Clin Epidemiol* 2010;63:513–23.
- [13] Atkins D, Best D, Briss PA, Eccles M, Falck-Ytter Y, Flottorp S, et al. Grading quality of evidence and strength of recommendations. *BMJ* 2004;328:1490.
- [14] Hartmann KE, McPheeters ML, Biller DH, Ward RM, McKoy JN, Jerome RN, et al. Treatment of overactive bladder in women. *Evid Rep Technol Assess (Full Rep)* 2009;(187):1–120. v.
- [15] Holden DJ, Harris R, Porterfield DS, Jonas DE, Morgan LC, Reuland D, et al. Enhancing the use and quality of colorectal cancer screening. *Evid Rep Technol Assess (Full Rep)* 2010;(190):1–195. v.
- [16] Clarke M, Hopewell S, Chalmers I. Clinical trials should begin and end with systematic reviews of relevant evidence: 12 years and waiting. *Lancet* 2010;376:20–1.
- [17] Clarke M, Hopewell S, Chalmers I. Reports of clinical trials should begin and end with up-to-date systematic reviews of other relevant evidence: a status report. *J R Soc Med* 2007;100:187–90.
- [18] Clarke M, Alderson P, Chalmers I. Discussion sections in reports of controlled trials published in general medical journals. *JAMA* 2002;287:2799–801.
- [19] Robinson KA, Goodman SN. A systematic examination of the citation of prior research in reports of randomized, controlled trials. *Ann Intern Med* 2011;154:50–5.
- [20] Brown P, Brunnhuber K, Chalkidou K, Chalmers I, Clarke M, Fenton M, et al. How to formulate research recommendations. *BMJ* 2006;333:804–6.