

## EDITORIAL

# Methods Guide for Authors of Systematic Reviews of Medical Tests: A Collaboration Between the Agency for Healthcare Research and Quality (AHRQ) and the Journal of General Internal Medicine

Gerald W. Smetana, MD<sup>1,2</sup>, Craig A. Umscheid, MD<sup>3,4</sup>, Stephanie Chang, MD<sup>5</sup>, and David B. Matchar, MD<sup>6,7</sup>

<sup>1</sup>Division of General Medicine and Primary Care, Beth Israel Deaconess Medical Center, Boston, MA, USA; <sup>2</sup>Harvard Medical School, Boston, MA, USA; <sup>3</sup>Center for Evidence-based Practice, University of Pennsylvania Health System, Philadelphia, PA, USA; <sup>4</sup>Department of Medicine, University of Pennsylvania Perelman School of Medicine, Philadelphia, PA, USA; <sup>5</sup>Center for Outcomes and Evidence, Agency for Healthcare Research and Quality, Rockville, MD, USA; <sup>6</sup>Health Services and Systems Research Program, Duke-National University of Singapore Graduate Medical School, Singapore, Singapore; <sup>7</sup>Department of Medicine, Duke University Medical Center, Durham, NC, USA.

J Gen Intern Med 27(Suppl 1):S1-3

DOI: 10.1007/s11606-012-2053-1

© The Author(s) 2012. This article is published with open access at Springerlink.com

Over the past 2 decades, systematic reviews have risen in number, quality, and impact. The sheer volume of work is remarkable. For example, the annual number of meta-analyses (a subset of systematic reviews) indexed by MEDLINE has grown from 273 in 1990 to 4,526 in 2010. Well-conceived and written systemic reviews serve many functions for stakeholders. First, they help clinicians apply evidence from the medical literature to patient care by critically appraising and summarizing what is often, for a given topic, a large amount of published clinical investigation. Systematic reviews are particularly useful when substantial practice variation exists, actual practice differs from published standards of care, clinical guidelines differ in their recommendations, and a large body of recent literature provides new insights that may modify recommendations from those of published guidelines.

Second, they can provide the basis for establishing and revising clinical guidelines as well as many quality assessment metrics applied to physicians, group practices, and hospitals. Third, they can inform future research agendas by defining important unresolved questions. Lastly, they draw attention to differences in findings across studies addressing similar research questions and propose a basis for the conflicting results. For all of these reasons, their impact can be substantial. For example, in one study of 170 journals in the fields of general internal medicine, family practice, nursing, and mental health, the average impact factor for systematic reviews was 26.5.<sup>1</sup> In contrast, the mean impact factor for the top 40 general medical journals is 7.4.

Guidelines to assist authors of systematic reviews in medicine have evolved. Published in 1999, the QUORUM

(Quality of Reporting of Meta-analyses) guideline for reporting systematic reviews<sup>2</sup> aimed to standardize and improve published reports of systematic reviews. Subsequent evolution of review methods, including increasingly rigorous assessments of the risk of bias and more frequent inclusion of observational data, prompted the development of an updated reporting tool, PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses), which was published in 2009.<sup>3,4</sup> PRISMA aims to standardize the reporting of systematic reviews; it offers less guidance to authors on the conduct and performance of such reviews. Guidelines also exist to assist authors in the conduct of reviews. Since 1994, the Cochrane Collaboration has published, and regularly updated, a detailed handbook for authors of systematic reviews.<sup>5</sup> This methods guide focuses primarily on reviews of randomized controlled trials of interventions. While developed for authors of Cochrane reviews, the handbook is freely available and has been a helpful resource for other authors of systematic reviews.

One mission of the Agency for Healthcare Research and Quality (AHRQ) is to solicit and publish systematic reviews (evidence reports and technology assessments) on topics to improve the clinical practice and delivery of health-care services. In 1997, AHRQ formed the Evidence-based Practice Center (EPC) Program to commission and oversee these reviews.<sup>6</sup> One of us (SC) directs the EPC program under the umbrella of the Center for Outcomes and Evidence, and one (DM) directed the Duke EPC. EPCs conduct reviews for use by a variety of groups, including national guideline groups, such as the US Preventive Services Task Force (USPSTF),<sup>7</sup> which uses reviews to inform screening and prevention guidelines, and payers, such as the Centers for Medicare and Medicaid Services. To improve the quality and consistency of EPC reports, the Agency has published methods guidance, developed by EPC authors (*Methods Guide for Effectiveness and Comparative Effectiveness*

Reviews).<sup>8</sup> This guidance, along with those of other groups such as the Cochrane Collaboration, the USPSTF,<sup>9</sup> and the Institute of Medicine<sup>10</sup>, forms the basis for a standards in the conduct of systematic reviews.

The editors of the AHRQ Methods Guide realized, however, that systematic reviews of medical tests pose unique challenges that are not adequately addressed in guidelines for authors of reviews of interventions or comparative efficacy. For example, the principle “outcome” of a study of a medical test is commonly a proxy or intermediate outcome. An illustration of this is ultrasound evaluation of the carotid arteries. The most common outcome in an evaluation of this test is the accuracy of the test in identifying clinically significant stenosis of the artery. Clinicians, while interested in this proxy outcome, would find more value in the ability of the test to predict clinically significant outcomes (such as 10-year risk of stroke or cardiovascular death) or the effect of subsequent carotid endarterectomy or stenting on stroke or death rates. A review of the operating characteristics of carotid ultrasound would optimally assess both the proxy result (as compared to a reference standard, in this case, invasive angiography) and the downstream result of testing on clinically significant outcomes.

Clinicians obtain medical tests for a number of non-overlapping reasons. These include screening, diagnosis, prognosis, and prediction of treatment response. In recognition of the unique challenges in conducting reviews of diagnostic tests, the Cochrane Collaboration has formed a working group specifically tasked with providing guidance in this arena. A draft version of their handbook, which is at present incomplete, has begun to address these challenges.<sup>11</sup>

AHRQ has also recognized the limitations of the *Methods Guide for Effectiveness and Comparative Effectiveness Reviews* (herein referred to as the *General Methods Guide*) when applied to studies of medical tests. In 2007, AHRQ convened an expert working meeting on the methodologic challenges in performing systematic reviews of medical tests. Four white papers were commissioned and presented on May 28–29, 2008.<sup>12</sup> The discussions from this meeting formed the basis for the Medical Test EPC workgroups, led by DM, then director of the Center for Clinical Health Policy Research and of the Duke EPC. Three EPC workgroups identified and addressed practical challenges in each step of conducting systematic reviews of medical tests (understanding the context, performing the review, and synthesizing the evidence). From these workgroups, EPC authors wrote nine draft papers providing guidance on steps for systematically reviewing medical test evidence that were either not covered in the existing *General Methods Guide* or that illustrated how to apply the *General Methods Guide* to medical test evaluation. An additional two workgroups addressed issues unique to genetic and prognostic tests. Each paper underwent extensive peer review by EPC investigators, external peer review, and public comment.

The Society of General Internal Medicine (SGIM) and the editorial leadership of the *Journal of General Internal Medicine* recognize that academic general internists share with AHRQ the desire to improve the quality of systematic reviews of medical tests through dissemination of methods guides to potential authors. AHRQ approached the *Journal's* editorial leadership and proposed a collaborative effort to review and publish this guide. The AHRQ Scientific Resource Center managed the peer and public review process through the usual Effective Health Care Program mechanisms. Two deputy editors from the *Journal* (GS and CU) reviewed the peer and public review comments, and author responses. All four of us reconciled any remaining issues and submitted a consensus letter to the corresponding author of each chapter with additional requests for revisions. In particular, we sought to expand the scope of the articles beyond EPC authors to provide relevant guidance to *all* authors of systematic reviews of medical tests. Likewise, we guided manuscript development so that the resulting chapters would be of value to readers of systematic reviews of medical tests who seek to determine the strengths and weaknesses of the review and its impact on clinical practice. We asked authors to identify potential differences between their chapters and the recommendations from the upcoming Cochrane handbook for systematic reviews of diagnostic test accuracy, and to comment on the basis for any disparities. The final versions of each chapter manuscript were submitted simultaneously to the *Journal* for typesetting and to AHRQ for public posting. AHRQ also developed online training modules for authors based on the content of these manuscripts.<sup>13</sup>

This supplement represents the final product of these efforts. The supplement covers 12 core aspects of the optimal conduct of systematic reviews of medical tests and serves as guidance for authors. However, each paper, or chapter, stands on its own. It is our sincere hope that EPC and non-EPC authors of systematic reviews, as well as other researchers and clinician readers, will find this collated Methods Guide for Systematic Reviews of Medical Tests to be helpful for the generation and appraisal of reviews, as well as the application of reviews to decision making about the use of specific medical tests in clinical practice.

---

**Open Access:** This article is distributed under the terms of the Creative Commons Attribution License which permits any use, distribution, and reproduction in any medium, provided the original author(s) and the source are credited.

**Corresponding Author:** Gerald W. Smetana, MD; Division of General Medicine and Primary Care, Beth Israel Deaconess Medical Center, Shapiro 621D, 330 Brookline Ave., Boston, MA 02215, USA (e-mail: gsmetana@bidmc.harvard.edu).

## REFERENCES

1. Montori VM, Wilczynski NL, Morgan D, Haynes RB. Systematic reviews: a cross-sectional study of location and citation counts. BMC Med. 2003;1:2.

2. **Moher D, Cook DJ, Eastwood S, Olkin I, Rennie D, Stroup DF.** Improving the quality of reports of meta-analyses of randomised controlled trials: the QUOROM statement. *Quality of Reporting of Meta-analyses*. *Lancet*. 1999;354:1896–900.
3. **Liberati A, Altman DG, Tetzlaff J, et al.** The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *Ann Intern Med*. 2009;151:W65–94.
4. PRISMA: Transparent reporting of systematic reviews and meta-analyses. Accessed March 2, 2012. Available at <http://www.prisma-statement.org/index.htm>.
5. **Higgins J, Green S, eds.** *Cochrane Handbook for Systematic Reviews of Interventions*. Version 5.1.0 (updated March 2011). The Cochrane Collaboration.
6. Agency for Healthcare Research and Quality. Evidence-based Practice Centers: synthesizing scientific evidence to improve quality and effectiveness in health care. Accessed March 2, 2012. Available at <http://www.ahrq.gov/clinic/epc/>.
7. Agency for Healthcare Research and Quality: U.S. Preventive Services Task Force (USPSTF). Accessed March 9, 2012. Available at <http://www.ahrq.gov/clinic/uspstfix.htm>.
8. Agency for Healthcare Research and Quality. *Methods Guide for Effectiveness and Comparative Effectiveness Reviews*. Rockville, MD: Agency for Healthcare Research and Quality. Accessed March 2, 2012. Available at: <http://www.effectivehealthcare.ahrq.gov/index.cfm/search-for-guides-reviews-and-reports/?pageaction=displayproduct&productid=318>.
9. Agency for Healthcare Research and Quality. U.S. Preventive Services Task Force Procedure Manual. AHRQ Publication No. 08-05118-EF. July 2008.
10. Institute of Medicine of the National Academies. *Finding What Works in Health Care: Standards for Systematic Reviews*; March 2011. Accessed March 9, 2012. Available at: <http://www.iom.edu/Reports/2011/Finding-What-Works-in-Health-Care-Standards-for-Systematic-Reviews.aspx>.
11. Cochrane Diagnostic Test Accuracy Working Group. *Handbook for DTA reviews*. Accessed March 9, 2012. Available at: <http://srdta.cochrane.org/handbook-dta-reviews>.
12. Agency for Healthcare Research and Quality. *Effective Health Care Program. Medical tests*. White paper series; Nov. 16, 2009. Accessed March 9, 2012. Available at: <http://www.effectivehealthcare.ahrq.gov/index.cfm/search-for-guides-reviews-and-reports/?pageaction=displayproduct&productid=350>.
13. Agency for Healthcare Research and Quality. *CME/CE Activities*. 2010. Accessed March 9, 2012. Available at: <http://www.effectivehealthcare.ahrq.gov/index.cfm/tools-and-resources/cmece-activities/#ahrq>.