VALUE IN HEALTH 17 (2014) 141–142

Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/jval

EDITORIAL

Tools for Health Care Decision Making: Observational Studies, Modeling Studies, and Network Meta-Analyses

Comparative effectiveness research incorporates study designs extending beyond the randomized controlled trial (RCT) [1], which is the touchstone for demonstrating therapeutic efficacy. The litany of RCT flaws is often recited: not real world, not real patients, not real settings, not available, not timely, and not affordable. Although not every RCT is well designed, there are established narratives and tools to aid health care decision makers in appraising and interpreting them. The Comparative Effectiveness Research Collaborative Initiative is a collective effort among the Academy of Managed Care Pharmacy, the National Pharmaceutical Council, and the International Society for Pharmacoeconomics and Outcomes Research (ISPOR) to provide tools for decision makers to assess studies that use nonexperimental methods important to comparative effectiveness research. To this end, ISPOR Good Practices Task Forces have developed tools to assess 1) prospective and retrospective observational studies, 2) modeling studies, and 3) network meta-analysis studies.

Questionnaires to Assess the Relevance and Credibility of a Study

The tools are straightforward in design and purpose. Each Task Force developed a questionnaire to assess the relevance and credibility of its assigned study design. The Task Forces wisely rejected linking the questionnaires to scorecards or checklists. Scoring systems encourage a false sense of precision [2]. Checklists tempt the user into "vote-counting," creating an implicit score. The purpose of the questionnaire is to guide the user to formulate an understanding of the strengths and weaknesses of each study reviewed.

The first-order consideration is whether a study is relevant to the issue at hand. The three questionnaires use common relevance criteria, derivative of the well-known patient population, interventions, comparators, outcomes, and settings (PICOTS) framework. The relevance criteria address population, interventions, outcomes, and context (settings and circumstances). Notably, the criteria ask whether important interventions and outcomes are missing. What is not reported is as important, and often more so, than what is reported. Selective reporting, focus on intermediate or short-term outcomes, bias in the comparator interventions, sparse attention to adverse effects—all are notorious threats to the evidence base for decision making [3].

If relevant, is the study credible? Domains to assess study credibility are specific to each study design. Credibility goes to the heart of the design, conduct, and reporting of a study. Are the results dependable for decision making? Are the results actionable? Of course, a single study is rarely definitive. The questionnaires presented here address individual studies, not the credibility of a body of evidence on an issue. The Comparative Effectiveness Research Collaborative Initiative has addressed appraisal and synthesis of a body of evidence elsewhere.

Assessing the credibility of observational and modeling studies as well as network meta-analyses can challenge even a familiar reader. The many questions concerning the credibility of observational studies no doubt reflect the difficulties in assessing causal treatment effects outside an RCT [4]. The questionnaire provides a guide through many of these difficulties; unfortunately, the implications for decision makers may not be as clear as one would like. Assessing the credibility of a modeling study follows a more linear and straightforward path. It is reflected in the questionnaire organization. Conceptually integrating the various domains is also somewhat intuitive. Last, network meta-analyses can be complex undertakings and understanding the credibility of results and conclusions complex as well. The questionnaire manages to skillfully highlight the important domains and offers considerable clarity. The arguably necessary detail is substantial and some notions may be difficult, but the document should provide the underpinnings needed by a studious decision maker to evaluate credibility.

Conflict of interest is a credibility domain in all three questionnaires. Questions for the conflict of interest domain are as follows: (a) Were there any potential conflicts of interest? and (b) If there were potential conflicts of interest, were steps taken to address these? Of concern, the three questionnaires differ in the guidance provided on how to apply the questions. The observational studies questionnaire is most informative on the first question. Users are advised to search public sources because the lack of disclosure of conflict should not be construed as absence of conflict. As to steps to address conflicts, the questionnaires do offer minimal advice; the modeling studies report is silent on the matter. Admittedly, how to address conflict of interest is a thorny issue [5]. ISPOR could provide leadership by creating a good practice Task Force on this topic.

Utility of the Questionnaires

Each Task Force conducted user testing, calculating multirater agreement for the credibility domain. Multirater agreement was less than 60% for the observational studies questionnaire, and ratings were not provided 15% to 36% of the time. The network meta-analysis questionnaire had an average agreement score of 72%, with a range of 42% to 91%. Network meta-analysis is a new method, and familiarity with this method is likely limited among

Source of financial support: The authors have no other financial relationships to disclose.

decision makers. The higher rate of agreement than reported for the more familiar observational studies suggests that the respondents were a group likely conversant with network analysis. There was insufficient response to user testing of the modeling tool to report the results.

Decision makers (and their organizations) vary in their capacity to conduct appraisal of study methodology. The Task Force reports endorse education of users and provide a Web-based training tool to aid users and promote uptake. An alternative, and also complementary, approach is to create a curated online inventory of completed questionnaires. Many users will be interested in similar questions and assessing the same articles. Experienced curators could review the questionnaires submitted, providing commentary and a learning resource to improve users' facility with the questionnaire. Some decision makers will find it more efficient to simply consult the repository than to conduct their own review. The inventory could also be indexed by the relevance criteria, so that decision makers would have access to a collection of reviewed articles relevant to the topic of interest.

Although such an online repository would require an investment of resources from ISPOR, the benefit to users would be substantial. The Agency for Healthcare Research and Quality has undertaken such a repository for systematic review data abstraction. The repository would pool effort and access to a collection of articles and topics larger than can be accomplished by any one organization alone. It is important for multiple reviewers to assess an article. But there is also a point of diminishing returns, and so further effort is an exercise in redundancy, not reproducibility.

Naomi Aronson, PhD,
Mark D. Grant, MD, PhD
Center for Clinical Effectiveness, Office of Clinical Affairs, Blue Cross Blue Shield Association, Chicago, IL, USA

1098-3015/$36.00 – see front matter
http://dx.doi.org/10.1016/j.jval.2014.02.001

REFERENCES

[1] Tunis SR, Benner J, McClellan M. Comparative effectiveness research: policy context, methods development and research infrastructure. Stat Med 2010;29:1963–76.
[2] Juni P, Witschi A, Bloch R, et al. The hazards of scoring the quality of clinical trials for meta-analysis. JAMA 1999;282:1054–60.
[3] Zarin DA, Tse T. Trust but verify: trial registration and determining fidelity to the protocol. Ann Intern Med 2013;159:65–7.
[4] Hernan MA, Robins JM. Estimating causal effects from epidemiological data. J Epidemiol Community Health 2006;60:578–86.
[5] Lo B, Field MJ, Institute of Medicine (U.S.), Committee on Conflict of Interest in Medical Research Education and Practice. Conflict of Interest in Medical Research, Education, and Practice. Washington, DC: National Academies Press, 2009.