A framework for best evidence approaches can improve the transparency of systematic reviews

Jonathan R. Treadwell[a,*], Sonal Singh[b], Ripple Talati[c], Melissa L. McPheeters[d], James T. Reston[a]

[a]ECRI Institute Evidence-Based Practice Center, 5200 Butler Pike, Plymouth Meeting, PA 19462, USA
[b]Johns Hopkins University Evidence-Based Practice Center, Department of Medicine, 624 N Broadway, Rm 680 B, Baltimore, MD 21287, USA
[c]University of Connecticut Evidence-Based Practice Center, 80 Seymour St, Hartford, CT 06102, USA
[d]Vanderbilt University Evidence-Based Practice Center, Vanderbilt University Medical Center, Suite 600, 2525 West End Ave., Nashville, TN 37203-1738, USA

Abstract

Objective: Systematic reviewers often use a ''best evidence'' approach to address the key questions, but what is meant by ''best'' is often unclear. The goal of this project was to create a decision framework for ''best evidence'' approaches to increase transparency in systematic reviews.

Study Design and Setting: The project was separated into three areas: 1) inclusion criteria, 2) evidence prioritization strategies, and 3) evaluative approaches. This commentary focuses only on the second task. The full report is available on the Effective Healthcare Web site of the Agency for Healthcare Research and Quality.

Results: The four identified strategies were as follows: 1) Use only the single best study; 2) Use the best set of studies; 3) Same as 2, but also consider whether the evidence permits a conclusion; and 4) Same as 3, but also consider the overall strength of the evidence. Simpler strategies (such as #1) are less likely to produce false conclusions, but are also more likely to yield insufficient evidence (possibly because of imprecise data).

Conclusion: Systematic reviewers routinely prioritize evidence in numerous ways. This document provides a conceptual construct to enhance the transparency of systematic reviewers' decisions. © 2012 Elsevier Inc. All rights reserved.

Keywords: Systematic review; Evidence synthesis; Best evidence; Meta-analysis; Evidence-based medicine; Evidence framework

1. Introduction

Systematic reviewers often use a ''best evidence'' approach to address the key questions in the reviews. What is meant by ''best,'' however, is often unclear. The phrase ''best evidence'' was used by Slavin in a 1995 article as an ''intelligent alternative'' to a meta-analysis of all available evidence on a given clinical question [1]. This approach was designed to allow exclusion of lower-quality studies (based on a priori criteria) if enough higher-quality studies are available. The underlying concept is *evidence prioritization* (i.e., prioritizing some studies over others), which is used by all systematic reviews.

In this commentary, ''best evidence'' refers to any strategy for prioritizing evidence. It can help ensure (but cannot

guarantee) that the review's conclusions will stand the test of time. However, reviewers face a variety of dilemmas regarding *how* to prioritize the evidence. Components such as risk of bias and applicability are themselves multifaceted, and the resulting complexity has spawned innumerable approaches for prioritizing evidence, with no organizing framework [2].

We recently authored a report that provides such a framework for defining the ''best evidence''; the full report appears on the Effective Healthcare Web site of the Agency for Healthcare Research and Quality (AHRQ) [3]. Essentially, the report addresses a reviewer's decisions about lowering the evidence threshold. Why might reviewers do this? How can it be done? The report, which is not intended to be prescriptive, can help reviewers improve the transparency of decisions made during the process of performing a systematic review. Such transparency serves the important function of enabling end users to assess a review's methodology and applicability [4].

* Corresponding author. Tel.: 610-825-6000; fax: 610-834-1275.
E-mail address: jtreadwell@ecri.org (J.R. Treadwell).

What is new?

- Systematic reviewers often use "best evidence" approaches, but reviews vary greatly in what this means.

- We created a framework of several "best evidence" approaches.

- This commentary discusses four strategies for prioritizing evidence.

- The strategies vary in the risk of inappropriate conclusions, as well as the risk of inappropriate *lack* of conclusions, and feasibility.

- Reviewers can use this framework to maximize transparency.

During a review, evidence can be prioritized at several stages, such as the search strategy, the inclusion criteria, the outcomes analyzed, and which studies will be pooled in a meta-analysis. Our report was organized around three tasks: 1) create a list of possible inclusion criteria, and for each criterion, create a list of factors that might affect a reviewer's decision to use it, 2) create a list of evidence prioritization strategies, and 3) list the ways in which evidence prioritization strategies might be formally evaluated. This commentary focuses only on the second task, evidence prioritization strategies.

2. Evidence prioritization strategies

After the set of included studies for a key question is determined, a reviewer must decide which studies comprise the "best evidence" set. We define this as the set of studies that will be assessed and/or analyzed in an attempt to answer the key question. Reaching this answer may or may not involve meta-analysis.

Studies not considered as part of the "best evidence" set, but still included, would be tabled but not used to inform conclusions. Some reviewers may choose to use all included studies in the attempt to draw evidence-based conclusions. If so, then the full list of included studies already defines the "best evidence" set.

Sometimes, however, the included studies are so variable in their risk of bias and/or applicability that some further prioritization is necessary. In this effort, several strategies can be used, and our workgroup outlined four of them (Table 1). We next discuss these, from the simplest to the most complex.

The simplest (strategy 1 in the table) is to take the single "best" study, and *using it alone*, determine what conclusions can be drawn. The definition of "best" would be based on a careful balance of both *risk of bias* and *applicability*. For example, this strategy might be used when evaluating an evidence base that contains a single, well-designed, well-conducted mega-trial in a real-world setting, and a few smaller trials that are more susceptible to bias and/or less applicable.

One example of strategy 1 involves the use of hormone replacement therapy for postmenopausal women. Some reviewers may rely solely on the Women's Health Initiative trial [5], based on its higher internal validity and generalizability than other trials. Other reviewers, of course, may believe it wrong to ignore all of the other trials.

The single best-study approach has the advantage of maximizing quality (i.e., minimizing risk of bias and maximizing applicability). However, it has three disadvantages: 1) the lack of scientific replication of findings, 2) the inability to determine consistency across studies (e.g., heterogeneity of effect sizes), and 3) the likelihood of low statistical power (if the study is not a mega-trial) precluding an answer to the key question (resulting in an evidence grade of insufficient).

A second strategy is to add studies that, relative to the single best study, are more susceptible to bias and/or less applicable. This permits measurement of cross-study

Table 1. Strategies for defining the "best evidence" set

Prioritization strategy	Which factors are *explicitly considered* when defining the best evidence set?					Comments
	Risk of bias	Applicability	Replication	Conclusiveness	Overall evidence strength	
1. Single best study	Yes	Yes	No	No	No	Easiest, but has the highest risk of being inconclusive
2. Best set of studies	Yes	Yes	Yes	No	No	Harder, and an increased risk of a false conclusion
3. Best set of studies, and also consider conclusiveness	Yes	Yes	Yes	Yes	No	Still harder, and further risk of a false conclusion
4. Best set of studies, and also consider conclusiveness and evidence strength	Yes	Yes	Yes	Yes	Yes	Still harder, and further risk of a false conclusion