

Chapter 3. Evidence-based Review Methodology

Definition and Scope

For this project the UCSF-Stanford Evidence-based Practice Center (EPC) defined a *patient safety practice* as “a type of process or structure whose application reduces the probability of adverse events resulting from exposure to the health care system across a range of conditions or procedures.” Examples of practices that meet this definition include computerized physician order entry (Chapter 6), thromboembolism prophylaxis in hospitalized patients (Chapter 31), strategies to reduce falls among hospitalized elders (Chapter 26), and novel education strategies such as the application of “crew resource management” to train operating room staff (Chapter 44). By contrast, practices that are disease-specific and/or directed at the underlying disease and its complications (eg, use of aspirin or beta-blockers to treat acute myocardial infarction) rather than complications of medical care are not included as “patient safety practices.” Further discussion of these issues can be found in Chapter 1.

In some cases, the distinction between patient safety and more general quality improvement strategies is difficult to discern. *Quality improvement practices* may also qualify as patient safety practices when the current level of quality is considered “unsafe,” but standards to measure safety are often variable, difficult to quantify, and change over time. For example, what constitutes an “adequate” or “safe” level of accuracy in electrocardiogram or radiograph interpretation? Practices to improve performance to at least the “adequate” threshold may reasonably be considered safety practices because they decrease the number of diagnostic errors of omission. On the other hand, we considered practices whose main intent is to improve performance above this threshold to be quality improvement practices. An example of the latter might be the use of computer algorithms to improve the sensitivity of screening mammography.¹

We generally included practices that involved acute hospital care or care at the interface between inpatient and outpatient settings. This focus reflects the fact that the majority of the safety literature relates to acute care and the belief that systems changes may be more effectively achieved in the more controlled environment of the hospital. However, practices that might be applicable in settings in addition to the hospital were not excluded from consideration. For example, the Report includes practices for preventing decubitus ulcers (Chapter 27) that could be applied in nursing homes as well as hospitals.

The EPC team received input regarding the scope of the Report from the Agency for Healthcare Research and Quality (AHRQ), which commissioned the report. In addition, the EPC team participated in the public meeting of the National Quality Forum (NQF) Safe Practices Committee on January 26, 2001. The NQF was formed in 1999 by consumer, purchaser, provider, health plan, and health service research organizations to create a national strategy for quality improvement. Members of the Safe Practices Committee collaborated with the EPC team to develop the scope of work that would eventually become this Report.

Organization by Domains

To facilitate identification and evaluation of potential patient safety practices, we divided the content for the project into different *domains*. Some cover “content areas” (eg, adverse drug events, nosocomial infections, and complications of surgery). Other domains involve identification of practices within broad (primarily “non-medical”) disciplines likely to contain promising approaches to improving patient safety (eg, information technology, human factors research, organizational theory). The domains were derived from a general reading of the

literature and were meant to be as inclusive as possible. The list underwent review for completeness by patient safety experts, clinician-researchers, AHRQ, and the NQF Safe Practices Committee. For each domain we selected a team of author/collaborators with expertise in the relevant subject matter and/or familiarity with the techniques of evidence-based review and technology appraisal. The authors, all of whom are affiliated with major academic centers around the United States, are listed on page 9-13.

Identification of Safety Practices for Evaluation

Search Strategy

The UCSF-Stanford Evidence-based Practice Center Coordinating Team (“The Editors”) provided general instructions (Table 3.1) to the teams of authors regarding search strategies for identifying safety practices for evaluation. As necessary, the Editors provided additional guidance and supplementary searches of the literature.

Table 3.1. Search strategy recommended by coordinating team to participating reviewers

- a) *Electronic bibliographic databases.* All searches must include systematic searches of MEDLINE and the Cochrane Library. For many topics it will be necessary to include other databases such as the Cumulative Index to Nursing & Allied Health (CINAHL), PsycLit (PsycINFO), the Institute for Scientific Information’s Science Citation Index Expanded, Social Sciences Citation Index, Arts & Humanities Citation Index, INSPEC (physics, electronics and computing), and ABI/INFORM (business, management, finance, and economics).
- b) *Hand-searches of bibliographies of retrieved articles and tables of contents of key journals.*
- c) *Grey literature.* For many topics it will necessary to review the “grey literature,” such as conference proceedings, institutional reports, doctoral theses, and manufacturers’ reports.
- d) *Consultation with experts or workers in the field.*

To meet the early dissemination date mandated by AHRQ, the Editors did not require authors to search for non-English language articles or to use EMBASE. These were not specifically excluded, however, and authoring teams could include non-English language articles that addressed important aspects of a topic if they had translation services at their disposal. The Editors did not make recommendations on limiting database searches based on publication date. For this project it was particularly important to identify systematic reviews related to patient safety topics. Published strategies for retrieving systematic reviews have used proprietary MEDLINE interfaces (eg, OVID, SilverPlatter) that are not uniformly available. Moreover, the performance characteristics of these search strategies is unknown.^{2,3} Therefore, these strategies were not explicitly recommended. The Editors provided authors with a search algorithm (available upon request) that uses PubMed, the freely available search interface from the National Library of Medicine, designed to retrieve systematic reviews with high sensitivity without overwhelming users with “false positive” hits.⁴

The Editors also performed independent searches of bibliographic databases and grey literature for selected topics.⁵ Concurrently, the EPC collaborated with NQF to solicit information about evidence-based practices from NQF members, and consulted with the project’s Advisory Panel (page 31), whose members provided additional literature to review.

Inclusion and Exclusion Criteria

The EPC established criteria for selecting which of the identified safety practices warranted evaluation. The criteria address the applicability of the practice across a range of conditions or procedures and the available evidence of the practices’ efficacy or effectiveness.

Table 3.2. Inclusion/Exclusion criteria for practices

<p><u>Inclusion Criteria</u></p> <ol style="list-style-type: none">1. The practice can be applied in the hospital setting or at the interface between inpatient and outpatient settings AND can be applied to a broad range of health care conditions or procedures.2. Evidence for the safety practice includes at least one study with a Level 3 or higher study design AND a Level 2 outcome measure. For practices not specifically related to diagnostic or therapeutic interventions, a Level 3 outcome measure is adequate. (See Table 3.3 for definition of “Levels”). <p><u>Exclusion Criterion</u></p> <ol style="list-style-type: none">1. No study of the practice meets the methodologic criteria above.
--

Practices that have only been studied outside the hospital setting or in patients with specific conditions or undergoing specific procedures were included if the authors and Editors agreed that the practices could reasonably be applied in the hospital setting and across a range of conditions or procedures. To increase the number of potentially promising safety practices adapted from outside the field of medicine, we included evidence from studies that used less rigorous measures of patient safety as long as the practices did not specifically relate to diagnostic or therapeutic interventions. These criteria facilitated the inclusion of areas such as teamwork training (Chapter 44) and methods to improve information transfer (Chapter 42).

Each practice’s level of evidence for efficacy or effectiveness was assessed in terms of study design (Table 3.3) and study outcomes (Table 3.4). The Editors created the following hierarchies by modifying existing frameworks for evaluating evidence⁶⁻⁸ and incorporating recommendations from numerous other sources relevant to evidence synthesis.⁹⁻²⁵

Table 3.3. Hierarchy of study designs*

Level 1. <u>Randomized controlled trials</u> – includes quasi-randomized processes such as alternate allocation
Level 2. <u>Non-randomized controlled trial</u> – a prospective (pre-planned) study, with predetermined eligibility criteria and outcome measures.
Level 3. <u>Observational studies with controls</u> – includes retrospective, interrupted time series (a change in trend attributable to the intervention), case-control studies, cohort studies with controls, and health services research that includes adjustment for likely confounding variables
Level 4. <u>Observational studies without controls</u> (eg, cohort studies without controls and case series)
* Systematic reviews and meta-analyses were assigned to the highest level study design included in the review, followed by an “A” (eg, a systematic review that included at least one randomized controlled trial was designated “Level 1A”)

Table 3.4. Hierarchy of outcome measures

Level 1. <u>Clinical outcomes</u> - morbidity, mortality, adverse events
Level 2. <u>Surrogate outcomes</u> - observed errors, intermediate outcomes (eg, laboratory results) with well-established connections to the clinical outcomes of interest (usually adverse events).
Level 3. <u>Other measurable variables with an indirect or unestablished connection to the target safety outcome</u> (eg, pre-test/post-test after an educational intervention, operator self-reports in different experimental situations)
Level 4. <u>No outcomes relevant to decreasing medical errors and/or adverse events</u> (eg, study with patient satisfaction as only measured outcome; article describes an approach to detecting errors but reports no measured outcomes)

Implicit in this hierarchy of outcome measures is that surrogate or intermediate outcomes (Level 2) have an established relationship to the clinical outcomes (Level 1) of interest.²⁶ Outcomes that are relevant to patient safety but have not been associated with morbidity or mortality were classified as Level 3.

Exceptions to EPC Criteria

Some safety practices did not meet the EPC inclusion criteria because of the paucity of evidence regarding efficacy or effectiveness, but were included in the Report because of their face validity (ie, an informed reader might reasonably expect them to be evaluated; see also Chapter 1). The reviews of these practices clearly identify the quality of evidence culled from medical and non-medical fields.

Evaluation of Safety Practices

For each practice, authors were instructed to research the literature for information on:

- prevalence of the problem targeted by the practice
- severity of the problem targeted by the practice
- the current utilization of the practice
- evidence on efficacy and/or effectiveness of the practice
- the practice's potential for harm
- data on cost if available
- implementation issues

These elements were incorporated into a template in an effort to create as much uniformity across chapters as possible, especially given the widely disparate subject matter and quality of evidence. Since the amount of material for each practice was expected to, and did, vary substantially, the Editors provided general guidance on what was expected for each element, with particular detail devoted to the protocol for searching and reporting evidence related to efficacy and/or effectiveness of the practice.

The protocol outlined the search, the threshold for study inclusion, the elements to abstract from studies, and guidance on reporting information from each study. Authors were asked to review articles from their search to identify practices, and retain those with the better study designs. More focused searches were performed depending on the topic. The threshold for study inclusion related directly to study design. Authors were asked to use their judgment in deciding whether the evidence was sufficient at a given level of study design or whether the evidence from the next level needed to be reviewed. At a minimum, the Editors suggested that there be at least 2 studies of adequate quality to justify excluding discussion of studies of lower level designs. Thus inclusion of 2 adequate clinical trials (Level 1 design) were necessary in order to exclude available evidence from prospective, non-randomized trials (Level 2) on the same topic.

The Editors provided instructions for abstracting each article that met the inclusion criteria based on study design. For each study, a required set of 10 abstraction elements (Table 3.5) was specified. Authors received a detailed explanation of each required abstraction element, as well as complete abstraction examples for the 3 types of study design (Levels 1A, 1, and 3; Level 2 was not included since the information collected was same as Level 1). Research teams were encouraged to abstract any additional elements relevant to the specific subject area.

Table 3.5. Ten Required Abstraction Elements

<ol style="list-style-type: none">1. Bibliographic information according to AMA Manual of Style: title, authors, date of publication, source2. Level of study design (eg, Level 1-3 for studies providing information for effectiveness; Level 4 if needed for relevant additional information) with descriptive material as follows: <i>Descriptors</i> <u><i>For Level 1 Systematic Reviews Only</i></u><ol style="list-style-type: none">(a) Identifiable description of methods indicating sources and methods of searching for articles.(b) Stated inclusion and exclusion criteria for articles: yes/no.(c) Scope of literature included in study. <u><i>For Level 1 or 2 Study Designs (Not Systematic Reviews)</i></u><ol style="list-style-type: none">(a) Blinding: blinded, blinded (unclear), or unblinded(b) Describe comparability of groups at baseline – ie, was distribution of potential confounders at baseline equal? If no, which confounders were not equal?(c) Loss to follow-up overall: percent of total study population lost to follow-up. <u><i>For Level 3 Study Design (Not Systematic Reviews)</i></u><ol style="list-style-type: none">(a) Description of study design (eg, case-control, interrupted time series).(b) Describe comparability of groups at baseline – ie, was distribution of potential confounders at baseline equal? If no, which confounders were not equal?(c) Analysis includes adjustment for potential confounders: yes/no. If yes, adjusted for what confounders?3. Description of intervention (as specific as possible)4. Description of study population(s) and setting(s)5. Level of relevant outcome measure(s) (eg, Levels 1-4)6. Description of relevant outcome measure(s)7. Main Results: effect size with confidence intervals8. Information on unintended adverse (or beneficial) effects of practice9. Information on cost of practice10. Information on implementation of practice (information that might be of use in whether to and/or how to implement the practice – eg, known barriers to implementation)

We present the salient elements of each included study (eg, study design, population/setting, intervention details, results) in text or tabular form. In addition, we asked authors to highlight weaknesses and biases of studies where the interpretation of the results might be substantially affected. Authors were not asked to formally synthesize or combine (eg, perform a meta-analysis) the evidence across studies for the Report.

Review Process

Authors submitted work to the Editors in 2 phases. In the first phase (“Identification of Safety Practices for Evaluation”), which was submitted approximately 6 weeks after authors were commissioned, authoring teams provided their search strategies, citations, and a preliminary list of patient safety practices to be reviewed. In the subsequent phase (“Evaluation of Safety Practices”), due approximately 12 weeks after commissioning, authors first submitted a draft chapter for each topic, completed abstraction forms, and—after iterative reviews and revisions—a final chapter.

Identification of Safety Practices for Evaluation

The Editors and the Advisory Panel reviewed the list of domains and practices to identify gaps in coverage. In addition, the Editors reviewed final author-submitted lists of excluded practices along with justifications for exclusion (eg, insufficient research design, insufficient outcomes, practice is unique to a single disease process). When there were differences in opinion as to whether a practice actually met the inclusion or exclusion criteria, the Editors made a final disposition after consulting with the author(s). The final practice list, in the form of a Table of Contents for the Report, was circulated to AHRQ and the NQF Safe Practices Committee for comment.

Evaluation of Safety Practices

Chapters were reviewed by the editorial team (The EPC Coordinating Team Editors and our Managing Editor) and queries were relayed to the authors, often requesting further refinement of the analysis or expansion of the results and conclusions. After all chapters were completed, the entire Report was edited to eliminate redundant material and ensure that the focus remained on the evidence regarding safety practices. Near the end of the review process, chapters were distributed to the Advisory Panel for comments, many of which were incorporated. Once the content was finalized, the Editors analyzed and ranked the practices using a methodology described in Chapter 56. The results of these summaries and rankings are presented in Part V of the Report.

References

1. Thurfjell E, Thurfjell MG, Egge E, Bjurstam N. Sensitivity and specificity of computer-assisted breast cancer detection in mammography screening. *Acta Radiol.* 1998;39:384-388.
2. Hunt DL, McKibbin KA. Locating and appraising systematic reviews *Ann Intern Med.* 1997;126:532-538.
3. Glanville J, Lefebvre C. Identifying systematic reviews: key resources. *ACP J Club.* 2000;132:A11-A12.
4. Shojania KG, Bero L. Taking advantage of the explosion of systematic reviews: an efficient MEDLINE search strategy. *Eff Clin Pract.* 2001;4: in press.
5. McAuley L, Pham B, Tugwell P, Moher D. Does the inclusion of grey literature influence estimates of intervention effectiveness reported in meta-analyses? *Lancet.* 2000;356:1228-1231.
6. Harris RP, Helfand M, Woolf SH, Lohr KN, Mulrow CD, Teutsch SM, et al. Current methods of the U.S. Preventive Services Task Force. A review of the process. *Am J Prev Med.* 2001;20:21-35.
7. Guyatt G, Schunemann H, Cook D, Jaeschke R, Pauker S, Bucher H. Grades of recommendation for antithrombotic agents. *Chest.* 2001;119:3S-7S.
8. Muir Gray JA, Haynes RB, Sackett DL, Cook DJ, Guyatt GH. Transferring evidence from research into practice: 3. Developing evidence-based clinical policy. *ACP J Club.* 1997;126:A14-A16.
9. Guyatt GH, Tugwell PX, Feeny DH, Drummond MF, Haynes RB. The role of before-after studies of therapeutic impact in the evaluation of diagnostic technologies. *J Chronic Dis.* 1986;39:295-304.
10. Davis CE. Generalizing from clinical trials. *Control Clin Trials.* 1994;15:11-4.
11. Bailey KR. Generalizing the results of randomized clinical trials. *Control Clin Trials.* 1994;15:15-23.
12. Oxman AD, Cook DJ, Guyatt GH. Users' guides to the medical literature. VI. How to use an overview. Evidence-Based Medicine Working Group. *JAMA.* 1994;272:1367-1371.
13. Levine M, Walter S, Lee H, Haines T, Holbrook A, Moyer V. Users' guides to the medical literature. IV. How to use an article about harm. Evidence-Based Medicine Working Group. *JAMA.* 1994;271:1615-1619.
14. Guyatt GH, Sackett DL, Sinclair JC, Hayward R, Cook DJ, Cook RJ. Users' guides to the medical literature. IX. A method for grading health care recommendations. Evidence-Based Medicine Working Group. *JAMA.* 1995;274:1800-1804.
15. Cook DJ, Guyatt GH, Laupacis A, Sackett DL, Goldberg RJ. Clinical recommendations using levels of evidence for antithrombotic agents. *Chest.* 1995;108:227S-230S.
16. Naylor CD, Guyatt GH. Users' guides to the medical literature. X. How to use an article reporting variations in the outcomes of health services. The Evidence-Based Medicine Working Group. *JAMA.* 1996;275:554-558.
17. Moher D, Jadad AR, Tugwell P. Assessing the quality of randomized controlled trials. Current issues and future directions. *Int J Technol Assess Health Care.* 1996;12:195-208.
18. Mulrow C, Langhorne P, Grimshaw J. Integrating heterogeneous pieces of evidence in systematic reviews. *Ann Intern Med.* 1997;127:989-995.
19. Jadad AR, Cook DJ, Browman GP. A guide to interpreting discordant systematic reviews. *CMAJ.* 1997;156:1411-1416.

20. Juni P, Witschi A, Bloch R, Egger M. The hazards of scoring the quality of clinical trials for meta-analysis. *JAMA*. 1999;282:1054-1060.
21. McKee M, Britton A, Black N, McPherson K, Sanderson C, Bain C. Methods in health services research. Interpreting the evidence: choosing between randomised and non-randomised studies. *BMJ*. 1999;319:312-315.
22. Bucher HC, Guyatt GH, Cook DJ, Holbrook A, McAlister FA. Users' guides to the medical literature: XIX. Applying clinical trial results. A. How to use an article measuring the effect of an intervention on surrogate end points. Evidence-Based Medicine Working Group. *JAMA*. 1999;282:771-778.
23. Concato J, Shah N, Horwitz RI. Randomized, controlled trials, observational studies, and the hierarchy of research designs. *N Engl J Med*. 2000;342:1887-1892.
24. Benson K, Hartz AJ. A comparison of observational studies and randomized, controlled trials. *N Engl J Med*. 2000;342:1878-1886.
25. Stroup DF, Berlin JA, Morton SC, Olkin I, Williamson GD, Rennie D, et al. Meta-analysis of observational studies in epidemiology: a proposal for reporting. Meta-analysis Of Observational Studies in Epidemiology (MOOSE) group. *JAMA*. 2000;283:2008-2012.
26. Gotzsche PC, Liberati A, Torri V, Rossetti L. Beware of surrogate outcome measures. *Int J Technol Assess Health Care*. 1996;12:238-246.

