
Robert H. Christenson,* Susan R. Snyder,2 Colleen S. Shaw,2 James H. Derzon,3 Robert S. Black,3 Diana Mass,4 Paul Epner,5 Alessandra M. Favoretto,3 and Edward B. Liebow3

OBJECTIVE: To develop methods for systematically reviewing evidence for identifying effective laboratory medicine (LM) practices associated with improved healthcare quality outcomes.

RELEVANCE: Although many evidence-evaluation systems have been developed, none are designed to include and rate healthcare quality improvement studies to identify evidence-based practices that improve patient safety and LM quality.

METHODS: Validated evidence-based medicine methods established by governmental agencies, the Guide to Community Preventive Services, and others were adapted for the LM field. Key methods modifications included (a) inclusion of quality improvement study designs; (b) mechanisms for inclusion of unpublished evidence, (c) combining of individual ratings of study quality, effect size, and relevance of outcome measures to evaluate consistency of practice evidence; and (d) deriving an overall strength rating to support evidence-based best practice recommendations. The methods follow the process steps of: ask; acquire; appraise; analyze; apply; and assess. Expert panels used the systematic evidence review results on practice effectiveness for improving healthcare quality outcomes consistent with the Institute of Medicine’s healthcare quality aims (safe, timely, effective, efficient, and patient-centered).

CONCLUSIONS: Adapting and developing methods from validated systems and applying them to systematically review and evaluate practices in LM by using published and unpublished studies is feasible. With these methods, evidence from quality improvement studies can be systematically synthesized and summarized to identify effective LM practices. Practical and scientifically validated demonstration of a positive impact on outcomes ensures that practitioners, policy makers, and decision makers at all levels have the evidence needed for improving healthcare quality and public health.

Nearly 100 000 deaths annually are attributable to medical error (1–3), prompting an increased emphasis on patient safety and quality improvement. Although more than 120 evidence-evaluation systems have been developed (4), none are designed to include observational quality improvement studies to identify evidence-based laboratory medicine (LM)5 practices. For this reason the Division of Laboratory Science and Standards (DLSS) of the CDC has supported development of the Laboratory Medicine Best Practices (LMBP) initiative, which is a systematic, transparent approach for evaluating evidence and identifying effective healthcare quality improvement practices. In this special report we describe the LMBP systematic evidence review methods for identifying evidence-based practices that are effective in improving healthcare quality and patient outcomes in a manner consistent with the 6 healthcare quality aims of the Institute of Medicine (IOM), to provide healthcare that is patient centered, safe, timely, effective, efficient, and equitable (1).

Practices are protocols, procedures, policies, techniques, processes, systems, standards, incentives, activities, and interventions that are used to provide healthcare to patients. The function of LM is to provide testing services that may have an impact on decisions...
regarding diagnosis, treatment, management, or prevention of disease, leading to actions intended to improve patient outcomes (5). Testing comprises the preanalytical (test selection, patient identification, specimen collection and identification), analytical (analyte measurement), and postanalytical (reporting of results and action) testing phases (6). Most errors occur in the preanalytical and postanalytical testing phases (7–10), and laboratory practices in these phases are the focus of LMBP efforts.

LMBP develops recommendations on the basis of careful review and synthesis of the evidence. This LMBP development process is transparent, scientifically sound, timely, and inclusive and is open to all relevant stakeholders and the public. The methods are designed so that given the same evidence, the review findings can be replicated by a different review team. Best-practice recommendations are issued by an independent recommending body that is not influenced by any particular faction, sponsoring agency, or political consideration, and for which all potential conflicts of interest have been fully disclosed. LMBP integrates with, rather than duplicates, existing efforts to conduct reviews for identifying and disseminating evidence-based practice recommendations.

Here we describe the LMBP systematic review process, which is aimed at the use of transparent, credible, evidence-based methods for evaluating practices and developing recommendations, as appropriate. A complete guide for the LMBP process is contained in the phase 3 final technical report (11) with examples of pilot LMBP efforts (also see Supplement A in the Data Supplement that accompanies the online version of this article at http://www.clinchem.org/content/vol57/issue7).

The A6 Cycle Method For Laboratory Medicine Best Practices

An overview of the LMBP systematic review process is shown in Figure 1. This process was derived, in part, from a standard strategy in evidence-based medical practice reviews (“Ask, Acquire, Appraise, Apply, and Assess”) (12, 13). An important addition for LMBP development is the inclusion of a specific “Analysis” step (14). Elements of validated evidence-based medicine methodologies established by the US Preventive Services Task Force (15, 16), Agency for Healthcare Research and Quality (17), Guide to Community Preventive Services (18), and Price and Christenson (19) were also included. Application of the LMBP methods requires coordinating the work of different groups, including an independent Workgroup, Review Team staff, and topic area Expert Panels. The Workgroup consists of experts in the LM field and other disciplines relevant to healthcare quality and evidence review methods; the Workgroup is convened by the CDC and has final decision-making responsibility for best practice recommendations. Review Team staff are trained in screening, abstracting, and rating studies for use as practice evidence. Expert Panelists are knowledgeable about the review-topic area, evidence review methods, and laboratory management. Expert Panelists are identified on the basis of their publication record as well as involvement and leadership in relevant organizations and initiatives. The Expert Panel applies the LMBP methods to review and evaluate the evidence synthe-
sized and ratings drafted by the Review Team to assess the strength of evidence for each practice, finalize evidence review and evaluation findings for each practice, and translate findings into draft evidence-based recommendations for consideration by the Workgroup.

**Topic Selection**

Candidate topics may be nominated by an individual or organization at www.futurelabmedicine.org. Responsibility for selecting and prioritizing the topics lies with the Workgroup. Topics selected must address a defined quality issue or problem in LM and satisfy the following 3 criteria:

- Evidence: availability of at least a modest body of evidence;
- Outcome measure(s): at least 1 relevant outcome consistent with IOM aims (1);
- Practices: at least 3 practices affecting performance or quality outcomes.

Following Workgroup acceptance, the nominators along with relevant Workgroup members and methodology experts collaborate to frame at least 1 focused review question.

**Asking The Question**

One or more focused review question(s) are formulated, a process aided by constructing an analytic framework based on information collected from preliminary review of published literature. (Online Supplement B shows the elements of an analytic framework.) In formulating focused questions and applying the analytic framework, the collaborative team must consider components of the PICO strategy (P, patient population; I, the intervention or practice; C, the comparator practice; and O, outcome) (12–16, 19). The analytic framework is revised to incorporate new knowledge during the review process.

**Acquiring the Evidence Base**

As with all systematic reviews, acquiring all the relevant evidence underpins the integrity of each subsequent step. The LMBP approach considers evidence from searches of published literature, and actively seeks evidence contributions from unpublished sources, as recommended by the PRISMA (Prevention Recovery Information System for Monitoring and Analysis) group (20), provided they meet the same study quality standards. Search terms derived from the review question, quality problem statement, and intervention and outcomes listed in the analytic framework are valuable for guiding evidence acquisition. Search strategies must involve a comprehensive literature search after a selected date relevant to current practice and technology. Consultation with Expert Panelists, authorities in other pertinent disciplines in the field, and Workgroup members is also useful for identifying relevant published and unpublished sources of evidence.

The LMBP initiative has made concerted efforts to access unpublished studies through contacts with experts and leaders among laboratory, hospital, and other healthcare consortia. Typically, these assessments are not termed “studies” or “research.” Nonetheless they may represent sufficiently rigorous and objective evaluations of high-quality data that constitute appropriate evidence for assessing practice effectiveness (21–23). Unpublished evidence is judged by use of the same criteria and standards as published studies.

For transparency, reports of LMBP systematic review results must list the search strategies, search terms, sources, and efforts used for locating both published and unpublished evidence. The number of studies identified by each strategy must also be specified in LMBP reports.

**Appraising Individual Studies for Inclusion**

The appraise step begins with an initial screening of published and unpublished search results from the acquire step. Screening is performed independently by 2 reviewers, who apply criteria to ensure that the practice described can be reproduced in comparable settings and has a potential impact on a quality-related outcome (LMBP inclusion/exclusion screening criteria are described in online Supplement C). For transparency, reporting of an LMBP evidence review must include a complete accounting of all potential evidence, specifying how many references and studies were identified during the search, how many met inclusion/exclusion screening criteria, how and why studies were eliminated during full-text review, and how many references and studies are included in the final review (20).

Full-text appraisal is accomplished by abstracting study information by use of an abstraction and appraisal tool (details provided in online Supplement C). To help control bias, 2 reviewers abstract each study independently and then compare their results. Discordances in study abstraction and ratings are reconciled through discussion. Unresolved discordances between reviewers require at least 1 additional reviewer to abstract the study and serve as the tiebreaker. Once detailed data are abstracted, an Evidence Summary Table is prepared for each study (see Table 1a), which summarizes bibliographic/source information and 4 dimensions related to study quality: study characteristics, practice description, outcome measures, and results/findings.
The rating of individual studies involves 3 sequential steps: (a) rating study quality; (b) rating the effect size(s), and (c) relevance to the review question ("direct," “less direct,” “indirect”).

Rating Individual Study Quality

As indicated in Table 1a, each quality dimension is assigned a score between 0 points and its maximum point value. The reviewers draft study quality ratings for each dimension along with justification for point deductions. This quality rating scheme was developed specifically by the LMBP Initiative, but these dimensions and their respective rating criteria were adapted from existing study quality rating instruments, checklists and theory (15, 16, 18, 22, 24–26). As such, they focus less on a study’s internal validity, and put greater weight on generalizability and potential for bias from sources outside the practice being tested. A "zero" rating on any dimension is sufficient to remove a study from consideration. For full details of the guidance and application of assigning quantitative values to the study quality rating criteria see the LMBP Guide to Rating Study Quality (online Supplement D) and LMBP Data Abstraction Codebook (online Supplement C).

Each dimension’s quality rating score from the Evidence Summary Table (Table 1a) is entered into the appropriate column of the Body of Evidence Table (Table 1b). The study dimension quality ratings are summed (maximum points, 10), and the quality rating of each individual study is rated as follows: good = 8–10; fair = 5–7; poor = 4 (Table 1b). Studies with “poor” quality ratings are excluded from a practice’s evidence base. After assignment of the individual study quality ratings, effect sizes for included studies are compiled in the right-most column of Table 1b.

Rating Individual Study Effect Size

Reviewers focus on the measure that most directly addresses the LMBP’s review question when abstracting the effect size element for the results/findings domain. The “relevance” rating addresses the directness of the relationship between the review question, the reported
outcome measure, and the relevant healthcare quality or safety outcome measure. A “direct” rating indicates a verifiable, direct, and proximate relationship between the problem or quality issue, the outcome measure reported, and the relevant IOM quality aim. A “less direct” rating indicates a less proximate (i.e., a longer causal chain) or greater reliance on inference. “indirect” indicates that the causal relationship requires logically supportable assumptions or is defined solely by associations with the quality aim. To facilitate comparisons when diverse outcome measures are used, study results are transformed to a standardized, common metric (an effect size). For dichotomous outcome measures expressed as rates or proportions (e.g., error rates, proportions of specimens meeting quality criteria, percent of results reported with a specified turn-around time), odds ratios are calculated. For continuous measures (mean and SD, such as time to an event, mean number of errors), the corresponding effect size is the standardized difference (Cohen’s $d$) (27).

The effect size for each study is classified as follows:

- Substantial: effect size is large enough to clearly support (or if adverse reject) practice implementation
- Moderate: effect size measured is large enough to support (or if adverse reject) practice implementation
- Minimal/none: effect size measured is statistically indistinguishable from the comparator practice or of no practical consequence

Analyzing the Evidence

The APPRAISE step focuses on individual studies, whereas the ANALYZE step involves the aggregate body of evidence. Whenever feasible, metaanalytic techniques should be used for summarizing the individual study effect sizes into an overall summary effect or grand mean estimate across all the studies. Individual study effect size estimates are weighted so that larger studies have more influence on the overall estimate of effect size and uncertainty (95% CI) for the practice (27, 28). It is useful to display the effect size for each individual study (and 95% CI) and the overall summary effect size from metaanalytic pooling of data using forest plots (Figure 2) (29). Forest plots are also helpful for representing the overall consistency of studies. Limitations of metaanalytic techniques are addressed in more detail in online Supplement E.

The Expert Panelists rate the overall effect size consistency and overall strength of the body of evidence (Table 2) using the compiled individual study ratings as well as any forest plots showing the individual study and overall effect size results. Following a precedent established by the Agency for Healthcare Research and Quality (17), overall consistency across individual studies for a given practice is treated as a dichotomous variable (i.e., “consistent” or “not consistent”) on the basis of similarity in reported effect sizes from studies for a given practice. A body of evidence for a given practice is considered “consistent” if the evidence is all in the same direction and within a reasonable narrow range of effect size magnitude. For the LMBP evaluation methods, “reasonableness” is determined by consensus expert judgment informed by the homogeneity and stability of the standardized study findings. The overall strength of a body of evidence rating is based on the total number of studies, their quality, and effect size ratings as indicated in Table 2.

The Expert Panel rates the overall strength of the body of evidence in support of the practice and it is categorized as High, Moderate, Suggestive, and Insufficient as defined (11). A brief explanation of the categories follows.

**HIGH**
Consistent evidence of substantial healthcare quality impact from studies without major limitations.

**MODERATE**
Some evidence of substantial healthcare quality impact from studies without major limitations; OR an adequate volume of consistent evidence of moderate healthcare quality impact from studies without major limitations.

**SUGGESTIVE**
Limited evidence of moderate healthcare quality impact from a small number of studies without major limitations; OR the quality of some studies’ design and/or conduct is limited.

**INSUFFICIENT**
Any estimate of an effect on healthcare quality impact is too uncertain.

These rating categories have their basis in the work of Guyatt et al. (24); they were modified to reflect both the quality of the evidence and effect size observed, rather than attempting to anticipate the impact of future potential evidence. The modified definitions for these categories are modeled after the US Preventive Services Task Force (16).

As shown in Table 2, the overall strength of the body of evidence ratings are derived from a minimum required number of studies after the Expert Panel has
For dichotomous outcome measures expressed as rates or proportions, Forest Plots of odds ratios are used to represent and compare diverse outcome measures from studies of the same best practice candidate. NS, not significant; Std diff, standard difference. For continuous outcome measures expressed as means/SDs, Forest Plots of standardized difference scores (e.g., Cohen's d) are used to represent and compare diverse outcome measures from studies of the same best practice candidate.
designated the various categorical ratings for individual study quality.

After the Expert Panel agrees on the overall strength of the body of evidence rating for a practice, an evidence-based LMBP draft recommendation based a strategy consistent with the GRADE group findings (24) is developed. This strategy was adapted for translating the overall strength of evidence rating into a recommendation that reflects the Expert Panel’s confidence that the practice(s) will do more good than harm, which also addresses actual and potential harms and benefits not assessed directly in analyzing and evaluating practice effectiveness. The LMBP rating categories for the overall strength of a body of evidence related to a potential best practice translates into the following recommendation rating categories. A practice is assigned 1 of 3 recommendations (www.futurelabmedicine.org).

RECOMMEND
The practice should be identified as a “best practice” for implementation in appropriate care settings, taking into account variations and applicability in implementation and/or care settings.

NO RECOMMENDATION FOR OR AGAINST
A potentially favorable impact on healthcare quality is not of sufficient size, or not sufficiently supported by evidence to indicate that it should be identified as a “best practice” for implementation in appropriate care settings.

RECOMMEND AGAINST
The practice should not be identified as a “best practice” for implementation because it is not likely to result in more good than harm.

The final stage is consideration by the Workgroup of the draft recommendations from the topic area Expert Panel. The Workgroup has responsibility for the final decision of the LMBP recommendations.

Applying the Results
Application of evidence is challenging. One study found that the proportion of recommended acute care
Clinical Chemistry 57:6 (2011)

provided to patients was only 53.5%, and stated that identified deficits in adherence to recommended processes pose serious threats to the health of the American public (30). Also, the IOM reported that on average it takes at least 17 years for new, effective medical research findings to become standard practice (2). The reasons for lack of adherence to evidence-based guidance are complicated and multifactorial (31–33). One important issue involves awareness of the evidence-based recommendations by stakeholder groups. Dissemination plans must engage key stakeholder groups including laboratory practitioners, laboratory professional organizations, clinicians, administrators, government regulatory groups, accrediting groups, policy makers, and payers. Dissemination and education related to LMBP methods and evidence reviews and recommendations should be accomplished through peer-reviewed publications, newsletters, press releases, and presentations at scientific and professional conferences. Access to LMBP recommendations, evidence summaries, and continuing education materials are available at www.futurelabmedicine.org.

Although LMBP recommendations are considered to have sufficient evidence of effectiveness, they may not address key questions commonly required by policy and decision-makers for implementation such as “Is it cost-effective and/or cost-saving?” and “Is it worth the cost?” To provide relevant information, one or more high-quality economic evaluation studies should be disseminated for recommended best practices that examine the cost implementation, and the medical and nonmedical costs, savings, and benefits from a societal perspective, which includes patients and their families, the healthcare system, and payers.

Potential harms and benefits associated with the practice must be considered. Potential benefits should include those that are not assessed directly in determining practice effectiveness. Examples of less direct benefits or harms may be impacts on patient, clinician, or healthcare worker satisfaction, ability to measure and monitor quality and process improvement, standardization of protocols, and other outcomes that affect patient safety and healthcare quality.

**Assess or Audit the Impact**

Assessing (validating) whether application of the LMBP is effective at improving LM services is important. The approach for internal and external validation of the LMBP systematic review process is outlined in Figure 1. Fundamental questions are: “Does the method produce a valid conclusion (Internal)?” and “Are the results generalizable across similar settings (External)?” Internal validation is performed for each topic conducted, e.g., online Supplement A. External validation requires comparison of the effect size as determined by metaanalytic procedures for institutions that adopt the recommended LMBP vs that predicted by the LMBP Systematic Review. This assessment can be accomplished by measuring/monitoring targeted outcomes to examine whether improvement has been realized. Results of these assessments should be forwarded to and collected by the LMBP to improve the base of evidence available for practice recommendations and to assess the generalizability of practice quality improvement results. The data needed for assessment does not necessarily require new studies, but is likely to be produced by local quality assurance/monitoring activities.

Evaluation, update, and revision of methods and results via new data collection are typically core functions of evidence review organizations. In the UK, the National Institute for Health and Clinical Excellence is an independent organization that provides technology appraisal guidance, develops clinical guidelines, and provides guidance for interventional procedures and public health. Formal auditing of the effectiveness of the efforts of the National Institute for Health and Clinical Excellence is a central activity (34).

Updating and refining of LMBP reviews and recommendations requires reapplying the A6 cycle by starting over with the ASK step and identifying new studies to use as evidence to evaluate the effectiveness and impact of quality improvement.

Intervention studies and economic evaluations can be designed to assess the impact of implementing LMBP findings and recommendations. Also, formation of registries to assess the impact of practice changes resulting from the LMBP is another assessment strategy.

**Conclusions**

Methods for systematically reviewing quality improvement practices in LM were developed from validated methods that have been broadly applied in other fields of medicine and public health. The methods are summarized by the A6 cycle: ASKING the review questions consistent with an analytic framework; ACQUIRING relevant evidence with particular efforts to identify unpublished studies; APPRAISING the evidence using screening criteria, full-text review, standardization, and evaluation/rating of individual study quality and effect size; ANALYZING the body of evidence and evaluating its sufficiency for developing evidence-based best practice recommendations; APPLYING the findings by dissemination and other promotional activities for implementing practice recommendations, and then ASSESSING the impact of practices and recommendations in the field. This methodology can be adopted to systematically summarize quality improvement evi-
idence and efficiently identify evidence-based LMBP for improving patient outcomes.

Author Contributions: All authors confirmed they have contributed to the intellectual content of this paper and have met the following 3 requirements: (a) significant contributions to the conception and design, acquisition of data, or analysis and interpretation of data; (b) drafting or revising the article for intellectual content; and (c) final approval of the published article.

Authors’ Disclosures or Potential Conflicts of Interest: Upon manuscript submission, all authors completed the Disclosures of Potential Conflict of Interest form. Potential conflicts of interest:

Employment or Leadership: None declared.

Consultant or Advisory Role: R.H. Christenson, Battelle Memorial Institute, Siemens Healthcare Diagnostics, Critical Care Diagnostics, and BG Medicine; D. Mass, Clinical Laboratory Management Association, Battelle Memorial Institute, and Brookline College; P. Epner, Battelle Memorial Institute.

Stock Ownership: None declared.

Honoraria: R.H. Christenson, Siemens Healthcare Diagnostics; D. Mass, COLA.

Research Funding: CDC (W911NF-07-D-0001/DO 0191/TCN 07235) to Battelle Centers for Public Health Research and Evaluation.


Role of Sponsor: The funding organizations played no role in the design of study, choice of enrolled patients, review and interpretation of data, or preparation or approval of manuscript.

Acknowledgments: Drs. Joe Boone and Roberta Carey provided substantial support and guidance, as did the LMBP Workgroup members and Expert Panelists. Additional assistance has been provided by Malaika Washington, Abrienne Patta (CDC), and Judy Berkowitz, Lisa John, Betsy Payn, and Shaynika Rose (Battelle).

References