Clinical Chemistry 57:6
816–825 (2011)

Special Report

Laboratory Medicine Best Practices: Systematic Evidence Review and Evaluation Methods for Quality Improvement

Robert H. Christenson,[1*] 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. Liebow[3]

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.

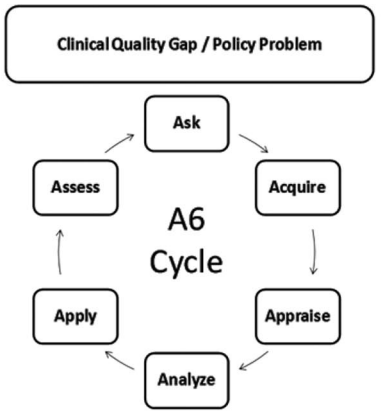
METHODS: Validated evidence-based medicine methods established by governmental agencies, the Guide to Community Preventive Services, and others were adapted for the LM field. Key methods modifications included (*a*) inclusion of quality improvement study designs; (*b*) mechanisms for inclusion of unpublished evidence, (*c*) combining of individual ratings of study quality, effect size, and relevance of outcome measures to evaluate consistency of practice evidence;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, equitable, 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from validated systems and applying them to systematically review and evaluate practices in LM by using published and unpublished studies is feasible.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)[6] 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

[1] University of Maryland School of Medicine, Baltimore, MD; [2] Centers for Disease Control and Prevention, Laboratory Research and Evaluation Branch, Division of Laboratory Science and Standards, Office of Surveillance, Epidemiology, and Laboratory Services, Atlanta, GA; [3] Battelle Centers for Public Health Research and Evaluation, North Seattle, WA; [4] Arizona State University, School of Life Sciences, Clinical Laboratory Sciences, Tempe, AZ; [5] Paul Epner LLC, 1501 Hinman Ave., Evanston, IL.
* Address correspondence to this author at: University of Maryland School of Medicine, 22 S. Greene St., Baltimore, MD 21201. Fax 410-218-6454; e-mail rchristenson@umm.edu.
Received October 4, 2010; accepted March 8, 2011.
Previously published online at DOI: 10.1373/clinchem.2010.157131
The findings and conclusions in this article are those of the authors and do not necessarily represent the views of the CDC.
[6] Nonstandard abbreviations: LM, laboratory medicine; DLSS, CDC Division of Laboratory Science and Standards, previously Division of Laboratory Systems; LMBP, Laboratory Medicine Best Practices; IOM, Institute of Medicine.



Validation of the LMBP Systematic Review Process

Internal Validation ("Appraise" and "Analyze" Steps)

- Double-coding of each study by staff reviewers using explicit instructions and objective criteria (Supplement C). A process of adjudication is specified for discrepant cases.
- The review process is designed to have examination and oversight review by an independent group, i.e, the Workgroup.
- Internal validation is conducted with each study, e. g, Supplement A.

External Validation ("Analyze," "Apply" and "Assess" Steps)

- Conduct LMBP systematic review on practices for a relevant laboratory topic, e.g., Supplement A.
- Utilize metaanalytic procedures to assess the consistency of findings and effect size across the included studies in accordance with LMBP methodology.
- Predict the range of outcomes for institutions that adopt the recommended LMBP.
- Assess the effect size for adopting the recommended LMBP, using metaanalytic procedures.
- Compare the observed effect size at the adopting institutions with the effect size predicted by the LMBP methodology.

Fig. 1. The Evidence-Based Practice Cycle adapted for laboratory medicine quality improvement and validation outline for the LMBP Systematic Review Process.

©Adapted with permission from the CDC on behalf of the LMBP initiative from CDC Laboratory Medicine Best Practices Team *(11)*. Supplements A and C can be found in the Data Supplement that accompanies the online version of this special report at http://www.clinchem.org/content/vol57/issue6. The Evidence-Based Practice Cycle and LMBP Systematic Review Process are transparent and replicable approaches to driving continuous improvement by identifying evidence-based best practices, with special emphasis in the preanalytic and postanalytic stages of the testing process, during which most preventable LM errors occur.

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 special report at http://www.clinchem.org/content/vol57/issue6).

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 prac-

tice 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 synthesized 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. (Supplement B in the online Data Supplement 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 Supplement C in the online Data Supplement). 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 Supplement C in the online Data Supplement). To help control bias, 2 reviewers abstract each study independently and then compare their results. Discordances in study abstraction and ratings are recon-

Table 1. Study characteristics abstracted for quality evaluation. ©CDC on behalf of the LMBP initiative, adapted with permission from CDC Laboratory Medicine Best Practices Team *(11)*.

a. Evidence summary table: abstracting and recording quality elements from full text review of each reference or study.[a]

Bibliographic information	Study characteristics	Practice description	Outcome measures	Results/findings
Author(s)	Design	Description	Description(s)	Type of findings
Year published/ submitted	Facility/setting	Duration	Recording method	Findings/effect size
Publication	Time period	Training		Statistical significance/test(s)
Author affiliations	Population/sample	Staff/other resources		Results/conclusion bias
Funding	Comparator Study bias	Cost		
Points for study quality domains	3 Points maximum	2 Points maximum	2 Points maximum	3 Points maximum

b. Collated information from the evidence summary table into part of the body of evidence table.

Citation from bibliography	Study characteristics[a]	Practice description	Outcome measures	Results/ finding	Total quality points[a]	Study quality rating (qualitative)[b]	Effect size rating (qualitative)[c]
Study 1	0–3	0–2	0–2	0–3	Maximum 10	Good, fair, or poor[b]	Substantial, moderate, etc.[c]
Study 2	0–3	0–2	0–2	0–3	Maximum 10	Good, fair, or poor[b]	Substantial, moderate, etc.[c]

[a] Range of values for each domain and total quality points from sum of the 4 domains; maximum 10-point scale.
[b] Conversion of total quality points to quality rating: G = good, 8–10 total quality points; F = fair, 5–7 total quality points; P = poor, ≤4 total quality points.
[c] Effect size for each study rated as: substantial, moderate, none/minimal, or adverse.

ciled 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*) rating of the study outcome measurement 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 grading process 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* (see Supplement D in the online Data Supplement) and *LMBP Data Abstraction Codebook* (see Supplement C in the online Data Supplement).

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 size ratings for included studies are compiled in the right-most column of Table 1b.

Rating Individual Study Effect Size

Reviewers focus on the study outcome 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 turnaround 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

The Expert Panel confirms the information abstracted into the practice Evidence Summary Table (Table 1a), and reaches consensus on the individual study quality and effect size ratings produced in the APPRAISE step. The consensus ratings are entered in the Body of Evidence Table (Table 1b).

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 of 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 Supplement E in the online Data Supplement.

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 effect size evidence is all in the same direction and within a reasonable narrow range of magnitude. For the LMBP evaluation methods, "reasonableness" is determined by consensus judgment of the topic area Expert Panel 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

Adequate volume of consistent evidence of substantial healthcare quality impact from studies without major limitations.

MODERATE

Some evidence of consistent 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.

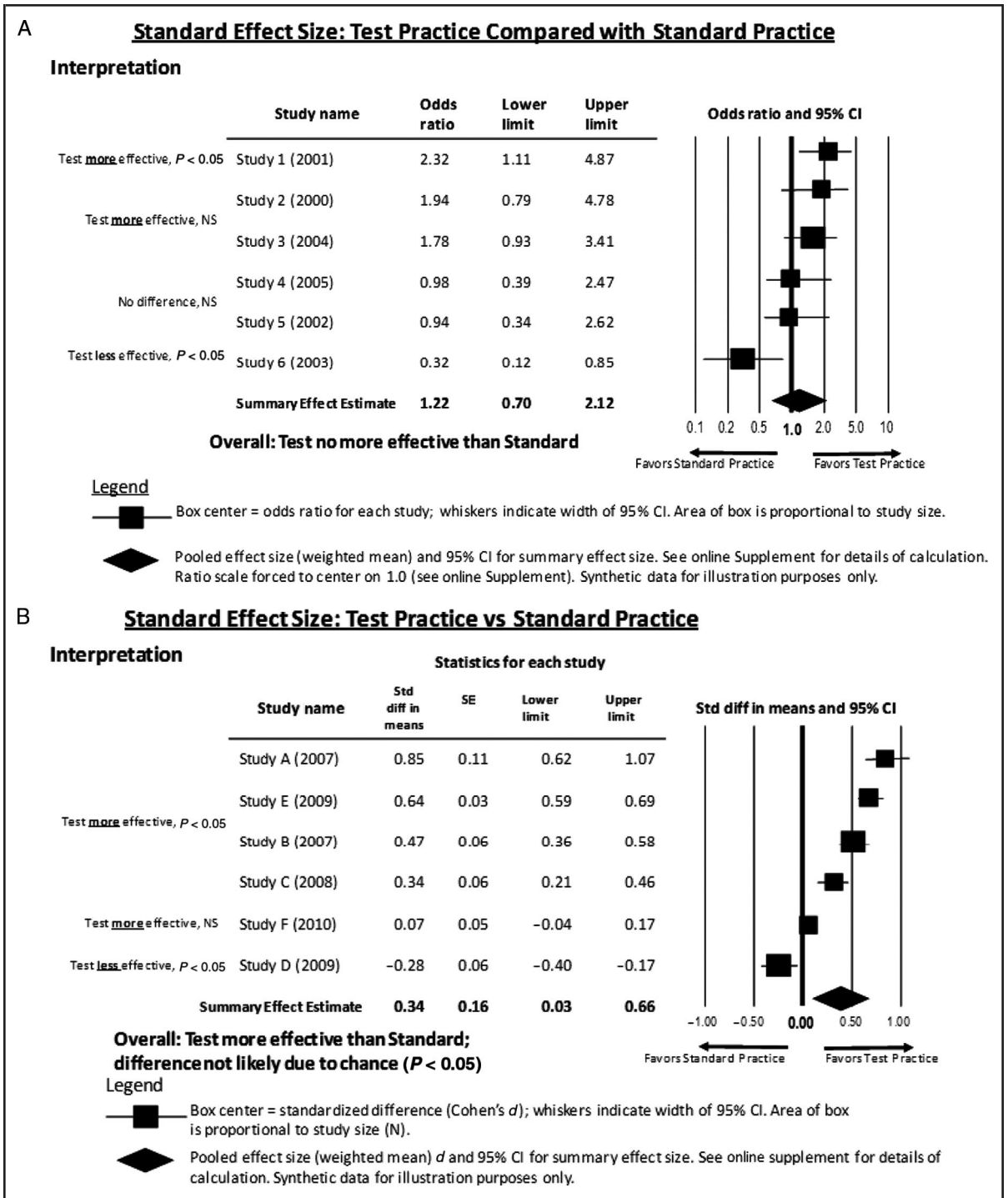


Fig. 2. (A), For dichotomous outcome measures expressed as rates or proportions, forest plots of odds ratios are used to represent and compare diverse outcome measures from studies of the same best practice candidate.

NS, not significant; Std diff, standard difference. (B), 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.

Table 2. Body of evidence tables.

a. Combining the Study Quality Rating from data abstraction [Table 4 of CDC Laboratory Medicine Best Practices Team *(11)*] with analysis for overall consistency and overall strength of body of evidence. ©CDC on behalf of the LMBP initiative, adapted with permission from CDC Laboratory Medicine Best Practices Team *(11)*.

	Study quality rating[a]								
Citation	Study characteristic	Practice description	Outcome measures	Results/ findings	Total	Quality rating	Effect size rating[c]	Overall consistency[d]	Overall strength of body of evidence[e]
Study 1 . . . study n	0–3	0–2	0–2	0–3	Maximum 10	G, F, or P[b]	See footnote[c]	Consistent OR not consistent[d]	High, moderate, suggestive OR insufficient[e]
Study 2									

b. Overall strength of body of evidence rating criteria. ©CDC on behalf of the LMBP initiative, adapted with permission from CDC Laboratory Medicine Best Practices Team *(11)*.

	Combined Evidence Minimum Criteria			
Strength Ratings	# Studies	Effect Size Rating		Quality Rating
High	≥3	Substantial	AND	Good
Moderate	≥2	Substantial	AND	Good
	OR ≥ 3	Moderate	AND	Good
Suggestive (Low)	≥1	Substantial	AND	Good
	OR ≥ 2	Moderate	AND	Good
	OR ≥ 3	Moderate	AND	Fair
Insufficient (Very Low)	All others			

[a] Study quality rating process presented in Table 1.
[b] G = good, 8–10 points total; F = fair, 5–7 points total; P = poor, ≤ 4 points total; see Table 4 for more detail.
[c] Effect size for each study rated as: substantial, moderate, or none/minimal.
[d] Overall consistency (among all studies) is a dichotomous variable (i.e., "consistent" OR "not consistent") based on similarity in reported effect sizes from studies included in a body of evidence for a given practice.
[e] See Panel b below for rating criteria for overall strength (among all studies).

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 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 on 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

High or moderate for improving healthcare quality (Table 2b). 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

Suggestive or insufficient (Table 2b). A potentially favorable impact on healthcare quality is not of sufficient size, or not sufficiently supported by evidence to indi-

cate that it should be identified as a "best practice" for implementation in appropriate care settings.

RECOMMEND AGAINST

High or moderate for adversely affecting healthcare quality (Table 2b). 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 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 of practice implementation, and the medical and nonmedical costs, savings, and benefits from a societal perspective, which includes the impacts on 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 evidence review conducted; e.g., see Supplement A in the online Data Supplement. 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 determine the effectiveness 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 evidence 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 Shyanika Rose (Battelle).

References

1. Institute of Medicine. To err is human: building a safer health system. Washington (DC): National Academy Press; 2000.
2. Institute of Medicine. Crossing the quality chasm: a new health care system for the 21st century. Washington (DC): National Academy Press; 2001.
3. Simmons J. Patient safety incidents showed little change in 2009 says HealthGrades. Healthleaders Media. http://www.healthleadersmedia.com/content/QUA-248824/Patient-Safety-Incidents-Showed-Little-Change-in-2009-Says-HealthGrades (Accessed July 2010).
4. West S, King V, Carey TS, Lohr KN, McKoy N, Sutton SF, Lux L. Systems to rate the strength of scientific evidence. Rockville (MD): Agency for Healthcare Research and Quality; 2002. Evidence report/technology assessment no. 47. Prepared by the Research Triangle Institute/University of North Carolina Evidence-Based Practice Center under contract no. 290-97-0011. AHRQ Publication No. 02-E016.
5. Chassin MR, Galvin R. The urgent need to improve health care quality: Institute of Medicine National Roundtable on Health Care Quality. JAMA 1998;280:1000–5.
6. Barr JT, Silver S. The total testing process and its implications for laboratory administration and education. Clin Lab Manage Rev 1994;8:526–42.
7. Bonini P, Plebani M, Ceriotti F, Rubboli F. Errors in laboratory medicine. Clin Chem 2002;48:691–8.
8. Carraro P, Plebani M. Errors in a STAT laboratory: types and frequencies 10 years later. Clin Chem 2007;53:1338–42.
9. Plebani M. Errors in clinical laboratories or errors in laboratory medicine. Clin Chem 2006;44:750–9.
10. Plebani M. Exploring the iceberg of errors in laboratory medicine. Clin Chim Acta 2009;404:16–23.
11. CDC Laboratory Medicine Best Practices Team. Laboratory medicine best practices: developing systematic evidence review and evaluation methods for quality improvement: Phase 3 final technical report. https://www.futurelabmedicine.org/pdfs/LMBP%20Executive%20Summary%20%20YR%203%20Final%20Report.pdf (Accessed April 2011).
12. Price CP, Lozar J, Christenson RH, eds. Applying evidence-based laboratory medicine: a step-by-step guide. Washington (DC): AACC; 2009.
13. Shaneyfelt T, Baum KD, Bell D, Feldstein D, Houston TK, Kaatz S, et al. Instruments for evaluating education in evidence-based practice: a systematic review. JAMA 2006;296:1116–27.
14. Kingsnorth J, O'Connell K, Guzzetta CE, Edens JC, Atabaki S, Mecherikunnel A, Brown K. Family presence during trauma activations and medical resuscitations in a pediatric emergency department: an evidence-based practice project. J Emerg Nurs 2010;36:115–21.
15. Harris RP, Helfand M, Woolf SH, Lohr KN, Mulrow CD, Teutsch SM, Atkins D. Current methods of the US Preventive Services Task Force: a review of the process. Am J Prev Med 2001;20:21–35.
16. US Preventive Services Task Force. GRADE definitions. Rockville (MD): Agency for Healthcare Research and Quality; 2008.
17. Agency for Healthcare Research and Quality. Methods reference guide for effectiveness and comparative effectiveness reviews. Version 1.0. Rockville (MD): Agency for Healthcare Research and Quality; 2007.
18. Zaza S, Briss P, Harris KW. The guide to community preventive services: what works to promote health? New York: Oxford University Press; 2005.
19. Price CP, Christenson RH, eds. Evidence-based laboratory medicine principles, practice and outcomes. Washington (DC): AACC; 2007.
20. Moher D, Liberati A, Tetzlaff J, Altman DG, and the PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. Ann Intern Med 2009;151:264–9.
21. Dreyer NA, Scheneeweiss S, McNeil BJ, Berger ML, Walker AM, Ollendorf DA, Gliklich RE. GRACE principles: recognizing high-quality observational studies of comparative effectiveness. Am J Manag Care 2010;16:467–71.
22. Ogrinc G, Mooney SE, Estrada C, Foster T, Goldmann D, Hall LW, Huizinga MM, et al. The SQUIRE (Standards for Quality Improvement Reporting Excellence) guidelines for quality improvement reporting: explanation and elaboration. Qual Saf Health Care 2008;17(Suppl 1):i13–32.
23. Horvath AR, Pewsner D. Systematic reviews in laboratory medicine: principles, processes and practical considerations. Clin Chim Acta 2004;342:23–39.
24. Guyatt GH, Oxman AD, Vist GE, Kunz R, Falck-Ytter Y, Alonso-Coello P, Schunemann HJ. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. BMJ 2008;336:924–6.
25. Higgins JPT, Green S, eds. Cochrane handbook for

systematic reviews of interventions 4.2.6. Updated September 2006. Chichester (UK): John Wiley & Sons; 2006.

26. von Elm E AD, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP. STROBE Initiative. Strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. J Clin Epidemiol 2008;61:344–9.
27. Borenstein M, Hedges LV, Higgins JPT, Rothstein HR. Introduction to meta-analysis. West Sussex (UK): Wiley & Sons; 2009.
28. Fleiss J. Measures of effect size for categorical data. In: Hedges LV, ed. The handbook of research synthesis. New York: Russell Sage Foundation; 1994. p 245–60.
29. Lewis S, Clarke M. Forest plots: trying to see the wood and the trees. BMJ 2001;322:1479–80.
30. McGlynn EA, Asch SM, Adams J, Keesey J, Hicks J, DeCristofaro A, Kerr EA. The quality of health care delivered to adults in the United States. N Engl J Med 2003;348:2635–45.
31. Turkleson C, Hughes J. Why aren't you doing evidence-based practice? AAOS Bull 2006;54:1–4.
32. Barnes MM. Why aren't community radiation oncologists adhering to evidence-based guidelines? Commun Oncol 2006;3:625–6.
33. Edelson E. Many doctors don't know blood pressure guidelines. ABC News/Health. http://abcnews.go.com/Health/Healthday/story?id=5844753&page=1 (Accessed July 2010).
34. National Institute for Health and Clinical Excellence. How to change practice: understand, identify and overcome barriers to change. London: National Institute for Health and Clinical Excellence; 2007.