

# Evidence-based Laboratory Medicine: Supporting Decision-Making

CHRISTOPHER P. PRICE

There is an implicit acceptance that an evidence-based culture underpins the practice of laboratory medicine, in part because it is perceived as the scientific foundation of medicine. However, several reviews of specific test procedures or technologies have shown that the evidence base is limited and in many cases flawed. One of the key deficiencies in the scientific literature on diagnostic tests often is the absence of an explicit statement of the clinical need, i.e., the clinical or operational question that the use of the test is seeking to answer. Several reviews of the literature on specific procedures have also demonstrated that the experimental methodology used is flawed with, in some cases, significant bias being introduced. Despite these limitations it is recognized that a more evidence-based approach will help in the education and training of health professionals, in the creation of a research agenda, in the production of guidelines, in the support of clinical decision-making, and in resource allocation. Furthermore, as knowledge and technologies continue to be developed, an evidence-based strategy will be critical to harnessing these developments.

© 2000 American Association for Clinical Chemistry

Evidence-based medicine has been hailed by some as a new paradigm for medical practice (1), whereas others claim that it is unscientific, with a heavy emphasis on a statistical and more managerial approach to decision-making that challenges the nature of clinical expertise and clinical decision-making (2). The proponents of an evidence-based approach point to the fact that clinical experience is a vital part of becoming a competent physician but that experience should be based on robust observation; robust in this context means the collection of data in a reproducible and unbiased way. Critics of the evidence-based approach point to the need for large randomized

controlled trials as the only means of meeting rigid criteria on acceptable evidence and the demise of the expert opinion (3). However, the rhetoric that surrounds the debate on evidence-based medicine needs to be put into context, recognizing that the burgeoning levels of knowledge and the increasingly multidisciplinary nature of healthcare are placing increasing demands on all practitioners (4).

Interestingly, to date evidence-based medicine appears to have had limited impact in the sphere of laboratory medicine. Furthermore, there are some data to suggest that adherence to criteria for the use of robust evidence in scientific papers on the use of diagnostic tests is poor (5). Laboratory medicine also provides some of the more overt examples of practice lacking a good foundation of evidence—perhaps the best examples being the variations seen in testing strategies between different hospitals for the same clinical presentations (6, 7). It is therefore hardly surprising that there are ardent critics of laboratory medicine and a considerable body of literature devoted to the inappropriate use of diagnostic tests (8–10). Perhaps the greatest challenge to laboratory medicine is the suggestion that diagnostic tests are not perceived to have a major impact on patient outcomes (11). Whereas most would consider this an extremely misguided viewpoint, it does indicate the degree of ignorance or misunderstanding that surrounds the value of diagnostic tests and poses one of the major challenges for today's laboratory professionals.

## Definitions and Concepts

### EVIDENCE-BASED MEDICINE

The generally accepted definition of evidence-based medicine is that given by Sackett et al. (12), namely “the conscientious, explicit and judicious use of current best evidence in making decisions about the care of patients”. This definition can readily be applied to laboratory medicine, recognizing (a) that a request for a diagnostic test represents part of a decision-making process; (b) the relevance to the clinical situation at hand; (c) the need for critical appraisal of evidence in terms of quality; and (d) the continuing evolution of evidence, particularly for new tests.

Department of Clinical Biochemistry, St. Bartholomew's and The Royal London School of Medicine & Dentistry, Turner Street, London E1 2AD, United Kingdom. Fax 44-20-7377-1544; e-mail c.p.price@mds.qmw.ac.uk.

Received March 2, 2000; accepted May 26, 2000.

#### DIAGNOSTIC TESTS

It has always been recognized that the use of a diagnostic test is an intervention (13). A diagnostic test should be requested only when a question is being posed and when there is evidence that the result will provide an answer to the question. There are several reasons why a physician will order a diagnostic test (14); however, the nature of the question being asked and the decision to be made often will depend on the clinical setting in which the patient is found.

#### OUTCOMES RESEARCH

Outcomes research, an important facet of evidence-based laboratory medicine, has been the subject of debate (13, 15, 16) over the past three decades. An outcome can be defined as the result of an intervention and may be a health outcome or an economic outcome (17, 18). This definition of an outcome may, in light of the limited perception of the contribution made by a diagnostic test (11), be a contributing factor to the paucity of good evidence on the effectiveness or benefit of diagnostic procedures in laboratory medicine.

The expectations for outcomes may be different for healthcare providers and patients (16). The patient is interested in receiving prompt and effective treatment, the relief of symptoms, and improvement in the quality of life. The service provider will also focus on the delivery of effective care, with the promptness of delivery varying between different countries but within the framework of optimum use of resources and minimization of long-term costs. Recognizing that any outcome will be a cascade of many synergistic decisions and actions, it may be more appropriate to focus greater attention on the use of diagnostic tests in the individual elements of the decision-making process. This may help to tease out the key constraints to delivering the desired patient outcomes while highlighting the value of the diagnostic test. In a recent editorial, Scott (19) stressed the importance of identifying a measurable outcome linked with the diagnostic procedure in question [point-of-care testing (POCT) in this case] (20).

#### DECISION-MAKING

Thus, the outcome of a diagnostic test can be considered as any part of the decision-making process that leads to an improved outcome. This approach was first described by Fryback and Thornbury (21), who set out the elements of clinical decision-making in relation to the efficacy of diagnostic tests. Thus, in terms of clinical benefit a test may improve the diagnostic process and/or the therapeutic strategy, and thus the overall health outcome. The outcome of a diagnostic test may be an operational or an economic benefit. Thus, in the example of the patient with chest pain, the first "questions" or decisions to be made relate to the recognition of cardiac pain and the urgency for referral. These early decisions are appropriate because

we know that early intervention improves overall patient outcome.

#### HEALTH TECHNOLOGY ASSESSMENT

Health technology assessment is a tool that examines systematically the consequences of the application of health technology to support decision-making in policy and practice. Technology encompasses drugs, devices, and procedures together with the organizational and support systems, therefore including diagnostic tests and differing modalities of delivery. However, the assessment focuses on the way in which the test or device is used rather than on whether it "works"; the latter is assumed and is the remit of another evaluation process (e.g., the Medical Devices Agency in the United Kingdom). The key principles of health technology assessment are outlined in Table 1.

Health technology assessment can be seen as the means by which evidence is developed to support decision-making, and it embraces much of the discipline of health services related or applied research. It may require original or primary research on the one hand, or a systematic review of established literature.

#### Context of the Evidence-based Culture

There are several reasons advanced for the development of a more evidence-based culture in medicine and the increasing commitment to outcomes research and health technology assessment. Not surprisingly these reflect the demands made on healthcare systems today (22), including some of the pressures listed in Table 2. All of these pressures translate into a common currency of cost pressures. It is worth exploring some of these pressures in a little more detail, focusing on issues related to implementation of new technology, decision-making, quality, and cost.

There is no doubt that healthcare professionals are expected to deal with an increasing burden of knowledge as well as having access to increasing amounts of new technology, both diagnostics and interventions. Earlier commentators have pointed to the numbers of scientific journals published each year and the limited time given to

**Table 1. Key principles of health technology assessment.**

Partnership	Recognize needs of purchasers, providers, and practitioners
Scientific credibility	Must meet high scientific standards
Independence	Free from influence of stakeholders
Accountability	Ultimately to the patient
Responsiveness	To perceived needs in a timely manner
Effectiveness	Audit to ensure process is effective
Visibility	Transparent to all parties
Accessibility	Available to all parties

**Table 2. Major pressures on healthcare systems that lead to the demand for an evidence-based approach to practice.**

Increased innovation	New technologies
Greater knowledge	New treatments, diagnostics, and rationales
Population changes	Ageing population, social changes
Increasing workload	More patient visits, more complex testing
Increasing spending	Salary costs, drug costs
Patient expectation	Greater knowledge (e.g., from the Internet), greater litigation

updating of knowledge (23). One of the strongest arguments for maintaining a health technology assessment program is the recognition that the process of evaluation and implementation of new technology is cumbersome, with long delays between clinical effectiveness being established and adoption into routine practice (24). There are examples in the literature, e.g., the implementation of thrombolytic therapy (25), that also demonstrate the importance of maintaining a continuous awareness of developments in the field to avoid unnecessary delays in the implementation of an effective new technology. It might also be argued that a lack of awareness of current literature leads to wastage of resources on additional evaluation projects when the answer is known. There are also examples in the literature where a new technology has been implemented ("technology creep") and subsequently considered to be ineffective or to lead to unnecessary interventions (26).

Quality in clinical practice has risen on the political agenda (22, 27), and the evidence-based culture has begun to play a greater role in policy-making on a global scale (28). The practical implications of the quality agenda have been a stronger focus on training and assessment of competence (29), on continuing education for maintenance of competence (30), and the introduction of clinical guidelines (31).

The rising cost of healthcare is also a topic high on the agenda of many governments and which is thought to contribute to the reduction in the number of hospital beds and the downsizing of laboratories (22, 32). It is also undoubtedly a stimulus for the literature on the appropriateness of laboratory testing and the demands for proper evaluation of new technology as a prerequisite for introduction into the healthcare system. However, it is also true that resource management is a major problem in the overall context of healthcare delivery, with little evidence that value for money plays a part in determining the investment in laboratory medicine for a given patient episode or disease state. For example, an examination of the economic impact of introducing a molecular-based test for *Chlamydia* and the influence on disease prevalence illustrates the complexity of resource allocation while

demonstrating the central role of good evidence in decision-making (33).

### How to Practice

Practicing evidence-based laboratory medicine has four dimensions: (a) identification of the question; (b) critical assessment of the best evidence available, embodying the principles of health technology assessment; (c) implementation of best practice; and (d) maintaining best practice; the latter embodies the principles of clinical audit (34). This is summarized in Fig. 1, which illustrates this as being a continuous process. There are many publications that provide guidelines on evaluation of the effectiveness of interventions (35, 36), although there are few that deal specifically with diagnostic tests. However, guidelines for the systematic review of diagnostic tests help by providing information on the critical appraisal of evidence (37–40).

#### DEFINE THE QUESTION

The starting point is the identification of the clinical question that is being asked; the ability, with the aid of the test result, to make a decision is therefore the outcome. There are many examples to illustrate this point, but it is not always recognized that a test can be used to make a "rule in" or "rule out" decision. Much of the literature in laboratory medicine has focused on "rule in" decisions, but "rule out" can often be an important step in a decision-making cascade. Silverstein and Boland (41) have suggested that when medical care costs are analyzed, the focus is directed to "high cost decisions"; they give the example of the decision to admit a patient. This may be a decision faced regularly by the primary care physician and may be addressed through a "rule in" or "rule out" question, depending on the actual diagnostic performance of the test. However, few diagnostic tests have been evaluated for the effectiveness in a "rule out" decision strategy. Examples where this approach has been proposed include urine testing for leukocyte esterase and nitrite, which can be effective in "ruling out" urinary tract infection but not "ruling in"; such a test can then be used

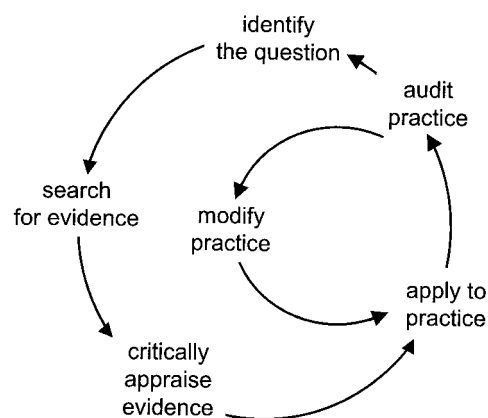


Fig. 1. The elements of evidence-based laboratory medicine.

to determine what samples are sent to a central laboratory for culture—with important operational and economic implications (42). Similarly it has been suggested that myoglobin could be used to “rule out” myocardial infarction, but not to “rule in” (43); in the past myoglobin has not been considered a useful marker of myocardial damage because of its lack of tissue specificity. This also illustrates the point that scientific reasoning may not always yield the correct interpretation in relation to clinical outcome—a point illustrated in the case of therapeutic interventions (44). The clear identification of a question can be particularly important when comparing the potential of a new test. Thus, the fact that serum cystatin C demonstrates a correlation with a reference clearance test superior to that of serum creatinine in itself does not prove that the test will offer a clinical benefit (45).

#### HIERARCHY OF EVIDENCE

Evidence on the performance of a diagnostic test can be considered in a hierarchy, all of the elements of which are important to making a decision (Fig. 2). This approach was first proposed for a diagnostic test by Fryback and Thornbury (21) and applied to diagnostic radiology (41).

#### TECHNICAL PERFORMANCE

The foundation of any evidence is technical performance, and this can have an important bearing on diagnostic performance. In addition to precision, accuracy, analytical range, and interferences, other pre-analytical factors such as biological variation and sample stability can influence the utility of a test. In general terms, laboratory professionals are extremely good at validating technical performance, although pre-analytical factors are less commonly documented (46). Pre-analytical factors can limit the benefits of a test in routine practice, e.g., the biological variation in the markers of bone resorption.

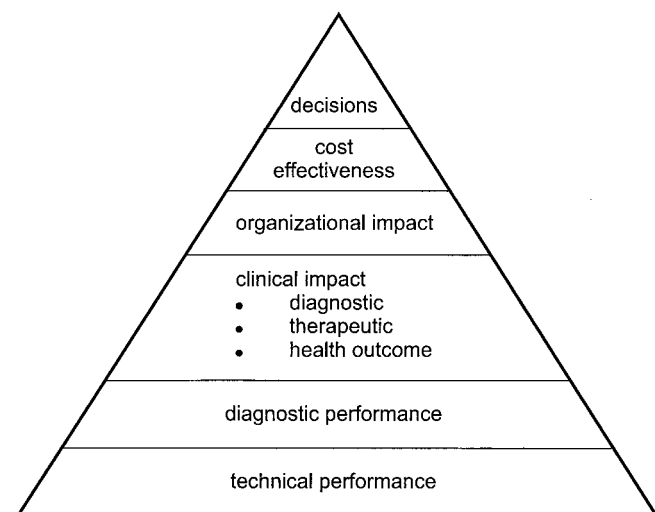


Fig. 2. Evidence of performance designed to facilitate decision-making.

#### DIAGNOSTIC PERFORMANCE

Diagnostic performance provides an assessment of the test in terms of the objective for using the test, namely the sensitivity, which defines the proportion of people who are correctly identified by the test as having the disease, and the specificity, which defines the proportion of people who are correctly identified by the test as not having the disease. Although these are parameters of any test irrespective of the population on which the test is used, the significance of the test is also determined by the prevalence of the condition in the population being studied. It has been suggested by Irwig et al. (38) and Moore (23) that the likelihood ratio combined with the pretest probability is a clearer way of identifying post test probability and thereby integrating the relevant information into a clinical decisions pathway. Batstone (47) has suggested that the number needed to diagnose (NND), derived from  $1/[\text{sensitivity} - (1 - \text{specificity})]$ , provides a useful comparison between tests and helps to encompass the financial implications in decision-making.

#### CLINICAL BENEFIT

However, it is the clinical impact or benefit of the test and the contribution to decision-making that provide the greatest challenge; the majority of evidence available in the literature on the use of diagnostic tests deals with technical and diagnostic performance. For example, Hobbs et al. (48), in a systematic review on POCT in general practice (primary care), found that few papers addressed clinical impact, the majority focusing on technical performance.

The clinical impact can be divided into the effect that use of a test or procedure will have (a) on the diagnostic strategy, i.e., compared with the use of other tests, in improving diagnostic performance; (b) on the therapeutic strategy, i.e., use of therapies, optimization of therapy, avoidance of harm, and so forth; and (c) on the health outcome. Thus, one can evaluate the impact of the detection of microalbuminuria in terms of (a) earlier detection of diabetic nephropathy and (b) better management of diabetes and co-morbid conditions, e.g., hypertension, with a view to (c) reducing the rate of development of renal failure. Some more examples of clinical impact are given in Table 3.

#### OPERATIONAL BENEFIT

The use of a diagnostic test may have an operational, as distinct from a clinical, impact; this often is considered an economic impact, which it may be because use of resources always resolves down to economic considerations. However, identification of an operational question may help to determine the optimal organizational aspects of a care strategy, e.g., disposition of staff and use of beds, as part of a wider economic analysis. Operational benefits may include reduced length of hospital stay, reduced staff time utilization, reduced utilization of estate (facilities), and reduced utilization of other resources. It may be

**Table 3. Examples of potential improved clinical benefit.**

Diagnostic strategy
"Rule out" of urinary tract infection
Triage of patients with acute chest pain
Screening tests
Therapeutic strategy
Improved compliance with therapy
Bone markers vs DEXA and bisphosphonate therapy in patients with osteoporosis
POCT for anticonvulsant drugs
POCT for HbA <sub>1c</sub>
Reduced risk of toxicity of therapy
Cyclosporin monitoring in transplant patients
Monitoring of lithium therapy
Clinical outcome
Reduced rate of onset of complications
Regular monitoring of glucose in patients with diabetes mellitus (including gestational diabetes)
Detection of microalbuminuria in patients with diabetes
Monitoring of proteinuria in patients at risk of preeclampsia
Screening tests
Testing for substance overdose, e.g., acetaminophen, salicylate

important to focus very specifically on the decisions that can follow from the identification of such a benefit, e.g., length of stay in relation to bed requirement. The decision-making will probably address two issues, namely bed utilization and clinical risk associated with early discharge. Rainey (49) considered length of stay as a medical outcome; however, this "definition" lacks the clarity of the decisions that need to be made.

#### ECONOMIC BENEFIT

The economic impact of the use of a diagnostic test and the broader issues of cost-benefit analysis are poorly understood tools in decision-making in healthcare. However, this is extremely important when the new test or procedure is more expensive than the existing modality of testing, as often is the case. It is also a wasted opportunity if the test brings real benefits to the patient and the healthcare organization. In these considerations, it is important not to focus solely on the test but on the complete patient episode or outcome (19, 50), determining where the investment is required and the gain achieved.

#### Evidence

Although the outcome of any evaluation of a test has to be recognition of the benefit, the starting point has to be identification of the question (or the clinical need). However, many studies fail to address the question, whereas others suffer from the use of poor methodology.

#### Systematic Review

Whereas the validation of a new test requires primary research, the accumulation of data from several studies in a systematic review can provide additional benefits. In-

deed, it could be argued that ongoing systematic review can provide a robust means of maintaining awareness of developments in a specific field. The benefits of a systematic review are summarized in Table 4 and are mainly concerned with handling of large amounts of data, minimizing the effects of bias introduced into individual studies, and increasing confidence in the result.

A systematic review of current evidence through the use of meta-analysis can identify methodological deficiencies as well as provide an up-to-date systematic review (51). The summary ROC curve can be used to combine studies, illustrating differences that may exist and determining their potential level of significance (52). Good examples of the use of summary ROC curve analysis for two clinical decisions are given in the reports by Guyatt et al. (53) and Olatidoye et al. (54). The work of Olatidoye et al. (54) demonstrates that the variability in diagnostic performance can be large and that this can present a dilemma to the evaluator. It also illustrates the need for multiple studies. Variability in performance may reflect differences in design, choice of patients, and clinical setting, all of which can lead to significant bias in the results (5, 55).

#### QUALITY OF EVIDENCE

Given that evidence informs a decision to be made, then it will be robust only if the evidence is of good quality. There have been many publications describing the quality of evidence that is acceptable (5, 40, 56–59) and identifying ways in which evidence can be gathered free of bias. In broad terms, bias can be introduced into evidence either from the choice of study population, the design of the experimental work, or the way in which the results are reported. The key priorities in designing a study to generate good quality evidence are summarized in Table 5.

The selection of the patient population for study has a critical bearing on the results produced and should be relevant to the question being asked. Thus the diagnostic accuracy of a test will be overestimated in a group of patients already known to have the disease when compared with a group of healthy individuals (59). Bias may also be introduced if the selection of patients for a trial is

**Table 4. Benefits to be gained from a systematic review of available evidence.**

Integrates information
Reduces data to manageable level
Identifies limitations in methodology
Limits effects of bias
Prevents proliferation of studies
Enhances confidence in overall result
Reduces delay between discovery and implementation
Establishes generalizability of data
Identifies causes of inconsistency
Identifies new research questions

**Table 5. Key study design criteria for achieving good quality evidence.**

Explicit identification of question
Appropriate choice of patients and setting
Application of reference standard to patients and controls
Use of robust measure of outcome
Blind comparison of method against reference or outcome
Exclusion of confounding variables
Description of methodology that can be reproduced

not randomized or if some patients are excluded from the trial; the latter may be the very patients on whom the quality of the decision hinges and may differentiate one test from another. Lijmer et al. (55) in a study of bias in 11 meta-analyses on 218 studies of diagnostic tests found that the largest effect on diagnostic accuracy occurred in case-control studies, suggesting that the mild cases, which are difficult to diagnose, were excluded, thereby causing overestimates of the sensitivity as well as specificity.

When assessing the diagnostic performance of a test, it is important to use a robust reference method. The reference method should be used in both the test and the control populations; this requirement cannot always be met, sometimes for ethical or cost reasons. However, it is recognized that failing to apply verification can lead to bias in the results. In some situations, to overcome this problem different reference methods have been applied to the test and control populations; again this can introduce bias, and it generated the second largest effect in the study by Lijmer et al. (55). As an example, in a study on the use of C-reactive protein (CRP) in the diagnosis of appendicitis, surgery and pathology were used as the reference for patients with high CRP, whereas verification was limited to clinical follow-up in the patients with low CRP concentrations (60).

It has also been suggested that blinding the evaluator of the new method to the results of the reference method will also improve the validity of the result. However, in the study by Lijmer et al. (55), it had little impact on the diagnostic accuracy of a test; conversely, in situations where the reference procedure is imperfect this effect may be increased. Valenstein (61), however, urged that rather than use an imperfect reference standard, one should focus on a more practical and measurable outcome; he used the example of the imperfect means of predicting the condition of a patient's myocardium, urging the use of a measurable outcome such as response to therapy or death in hospital. An imperfect reference procedure raises the prospect of interoperator variation, which can also contribute to bias (62). A good example is the diagnosis of myocardial infarction, where the evaluation of new biochemical markers introduces the bias associated with the use of other markers in the diagnostic triad but also recognizes the fact that there is interoperator variation in the ultimate reference procedure—the autopsy (63).

It is evident from the literature that study design

related to outcomes has been investigated more in relation to the use of a pharmaceutical intervention than a diagnostic intervention. Thus, Moore and Fingerova (64) have identified other characteristics of study design that will introduce bias to the results. The ideal approach, however, is identical whether it is a diagnostic or a therapeutic intervention and whether it is an assessment of diagnostic performance or outcome, and that is the use of prospective blind comparison of test and reference procedure in a consecutive series of patients from a relevant clinical population (60).

Studies designed to assess clinical, operational, and/or economic outcomes require a clear definition of the outcome measure. In certain clinical studies, this may require the use of a surrogate because the true outcome can only be assessed over several decades; examples include the use of hemoglobin (Hb)<sub>A1C</sub> as a surrogate for normoglycemic control and the DEXA scan for normal bone mineral density. An alternate outcome may be the "avoidance of disease" as was used in the Diabetes Control and Complications Trial (65); the approach currently underpins our assessment of wellness. It is also important to be aware of confounding factors that might also influence the outcome measure; thus in the study by Kendall et al. (66) on the use of POCT in the emergency room, the "time to result" was reduced, but the "length of stay" in the emergency room (potentially a very relevant outcome measure) was not influenced. The demonstration of reduced time to result, however, is still valid and points to another step in the process as being the impediment to reducing the length of stay.

It is also clear that publication of results also introduces bias; there is a greater tendency to publication of positive findings. Easterbrook et al. (67) showed in a study of a

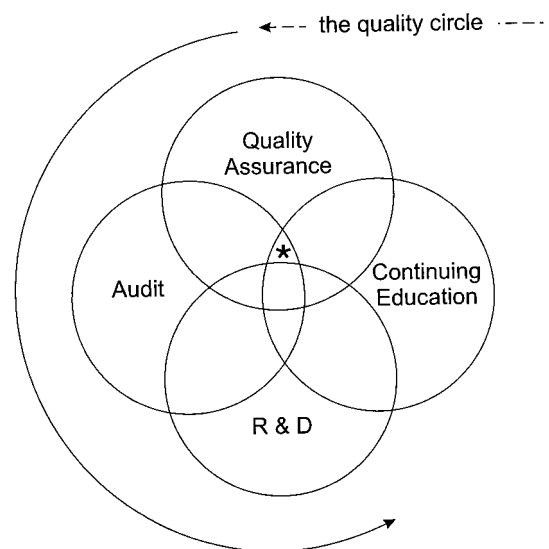


Fig. 3. Evidence-based laboratory medicine at the core (\*) of continuous quality improvement.

R & D, research and development.

large number of submissions to a local ethics committee that eventual publication was more prevalent in studies where there were significant findings. Dickersin et al. (68) also found that there was an association between significant results and publication, the bias originating with the authors rather than journal editors. Chalmers et al. (69) have defined three stages of publication: (a) organizing and undertaking of the research; (b) acceptance or rejection of a manuscript depending on the presentation of positive or negative findings; and (c) the bias that may result from interpretation, reviews, and meta-analyses. The authors also make recommendations on how publication bias can be minimized.

### Application

Reference has already been made to the fact that there often are delays in the implementation of new technologies despite the availability of good evidence, whereas on the other hand, technologies are implemented in the absence of evidence. It is acknowledged that even when there are health technology assessment programs set up to resolve some of these problems, the implementation of findings is still one of the major problems in the effectiveness of the programs (70). There may be several reasons for this, including (a) the absence of formal continuing education programs, and (b) the mechanisms for allocating resources. It is possible that initiatives that have been set up to identify and implement good practice, such as the National Institute for Clinical Excellence (71) coupled with the development of clinical governance (27) in the

UK, will improve implementation because both initiatives are founded on the culture of evidence-based practice.

### CLINICAL AUDIT

The establishment of a new practice, as well as established practice, should always be subject to regular audit (72); this underpins the commitment to maintenance of good practice, one of the important tenets of clinical governance. The audit will assess whether the new technology has been implemented satisfactorily and whether the outcomes found bear out the findings of the original research. The outcome of the audit may identify the need to modify practice or may lead to the identification of a new research question. Experience has shown that auditing established practice as, for example, a means of controlling demand for a laboratory service, can identify unmet clinical needs as well as abuse and inappropriate use of laboratory services. Thus, evidence-based laboratory medicine in all its facets is the foundation of a continuous quality improvement program (Fig. 3).

### Implications of an Evidence-based Culture

The foregoing discussion has been focused primarily on decision-making, but it is quite clear that the practice of evidence-based laboratory medicine can have an important influence on many aspects of professional practice.

### TRAINING AND MAINTAINING PERFORMANCE

It has been suggested that the principles of evidence-based medicine will provide a better foundation for

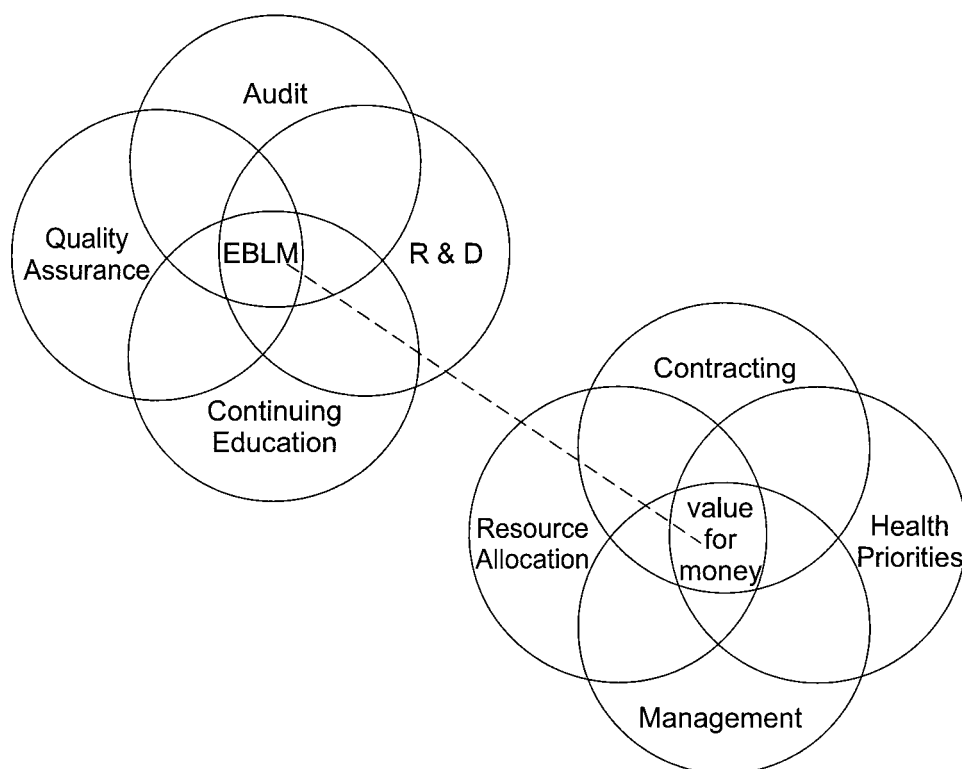


Fig. 4. Evidence-based laboratory medicine and the achievement of value for money.

EBLM, evidence-based laboratory medicine; R & D, research and development.

training because it focuses more on the evidence to support the type of decision-making required in clinical practice (73). It is also clear that performance deteriorates with age (74) and that traditional means of continuing education are less effective than an evidence-based approach (75). Thus, through participation in activities to maintain the evidence database and participation in audit activities, professional performance can be maintained and kept up to date.

#### RESEARCH AND DEVELOPMENT AGENDA

The continuous process of practice review, the generation of new knowledge, and the availability of new technology combine to generate a powerful development agenda. The use of evidence-based practice guidelines will ensure that an effective program can be developed to meet patients' needs in a timely and cost-effective fashion.

#### DECISION-MAKING

To establish the role of laboratory medicine in clinical decision-making, it is important to develop the type of evidence that focuses on these decisions. This requires an explicit recognition that medical research can be delineated into that which creates basic knowledge and that which is associated with the application of that knowledge. Applied research itself can then be differentiated into the application of knowledge in the development of diagnostic procedures and therapeutic interventions and that which focuses on the use of such innovation in decision-making (76). Evidence can then support decision-making in relation to diagnosis and therapeutic intervention, together with operational issues, including the appropriate utilization of resources.

#### VALUE FOR MONEY

The increasing cost of healthcare is one of the major pressures affecting both purchasers and providers of care. It is also evident that on the one hand, there is a perception that the cost of laboratory medicine is high, whereas on the other there is limited perception of true value for money—both in terms of the true cost and the framework in which value can be judged. The most obvious is the focus on the debate between the cost and value of POCT; invariably, and not unexpectedly, the cost of a point-of-care test is greater than its central laboratory counterpart, with the clinical and operational benefits accruing to other sections of the provider system (77). Thus, the value will only be appreciated outside of the laboratory, thereby requiring a wider perspective or review of value than the confines of the laboratory service.

#### QUALITY

An alternative description of evidence-based medicine is a commitment to life-long problem-based learning (73). A commitment to the activities outlined in Fig. 4 working within a framework as outlined in Fig. 1 will ensure that a high quality of service is maintained; these attributes are

embodied in the principles of laboratory accreditation (78), professional self-regulation (79), and ultimately clinical governance (27), representing commitment to provision of the highest quality of service to the patient.

#### Conclusions

Application of the principles of evidence-based medicine to laboratory medicine highlights the importance of establishing the role of diagnostic procedures in clinical decision-making. The discipline is crucial to creating and delivering the research and development agenda in the laboratory while also providing a foundation for the training of laboratory professionals. The continuing application of an evidence-based approach to practice will also meet the quality expectations of patients.

#### References

1. Guyatt G. Evidence-based medicine. A new approach to teaching the practice of medicine. *JAMA* 1992;268:2420–5.
2. Charlton BG, Miles A. The rise and fall of EBM. *Q J Med* 1998;91:371–4.
3. Antman EM, Lau J, Kupelnick B, Mosteller F, Chalmers TC. A comparison of results of meta-analyses of randomized control trials and recommendations of clinical experts. Treatments for myocardial infarction. *JAMA* 1992;268:240–8.
4. Oxman AD, Guyatt GH, Singer J, Goldsmith CH, Hutchison BG, Milner RA, Streiner DL. Agreement among reviewers of review articles. *J Clin Epidemiol* 1991;44:91–8.
5. Reid CK, Lachs MS, Feinstein AR. Use of methodological standards in diagnostic test research. Getting better but still not good. *JAMA* 1995;274:645–51.
6. Broughton PMG, Worthington DJ. Laboratories respond differently to the same clinical request. *Ann Clin Biochem* 1989;26:119–21.
7. Barth JH, Seth J, Howlett TA, Freedman DB. A survey of endocrine function testing by clinical biochemistry laboratories in the UK. *Ann Clin Biochem* 1995;32:442–9.
8. van Walraven C, Naylor CD. Do we know what inappropriate laboratory utilization is? A systematic review of laboratory clinical audits. *JAMA* 1998;280:550–8.
9. Perraro F, Rossi P, Liva C, Bulfoni A, Ganzini G, Giustinelli A, et al. Inappropriate emergency test ordering in a general hospital: preliminary reports. *Qual Assur Health Care* 1992;4:77–81.
10. Bareford D, Hayling A. Inappropriate use of laboratory services: long term combined approach to modify request patterns. *Br Med J* 1990;301:1305–7.
11. WD. Outcomes research in diagnostics: not yet. *In Vivo* 1995;Jan: 11–7.
12. Sackett DL, Rosenberg WMC, Gray JAM, Haynes RB, Richardson WS. Evidence-based medicine: what it is and what it isn't. *Br Med J* 1996;312:71–2.
13. Lundberg GD. The need for an outcomes research agenda for clinical laboratory testing. *JAMA* 1998;280:565–6.
14. Lundberg GD. How clinicians should use the diagnostic laboratory in a changing medical world. *Clin Chim Acta* 1999;280:3–11.
15. Witte DL. Measuring outcomes: why now? *Clin Chem* 1995;41: 775–80.
16. Wong ET. Improving laboratory testing: can we get physicians to focus on outcome? *Clin Chem* 1995;41:1241–7.
17. Bissell MG, ed. Laboratory-related measures of patient outcomes: an introduction. Washington: AACC Press, 2000:1–193.
18. Vance RP, Elevitch FR. College of American Pathologists Foundation Conference IVIII on patient-centered pathology practice—



- outcomes and accountability. *Arch Pathol Lab Med* 1997;121:1151–62.
19. Scott MG. Faster is better—it's rarely that simple [Editorial]. *Clin Chem* 2000;46:441–2.
  20. Nichols JH, Kickler TS, Dyer KL, Humbertson SK, Cooper PC, Maughan WL, et al. Clinical outcomes of point-of-care testing in the interventional radiology and invasive cardiology setting. *Clin Chem* 2000;46:543–50.
  21. Fryback FG, Thornbury JR. The efficacy of diagnostic imaging. *Med Decis Making* 1991;11:88–94.
  22. Price CP, Barnes IC. Laboratory medicine in the United Kingdom: 1948–1998 and beyond. *Clin Chim Acta* 2000;290:5–36.
  23. Moore RA. Evidence-based clinical biochemistry. *Ann Clin Biochem* 1997;34:3–7.
  24. Granados A, Jonsson E, Banta HD, Bero L, Bonair A, Cochet C, et al. Eur-Assess project subgroup report on dissemination and impact. *Int J Technol Assess Health Care* 1997;13:220–86.
  25. Lau J, Antman EM, Jimenez-Silva J, Kupelnick B, Mosteller F, Chalmers TC. Cumulative meta-analysis of therapeutic trials for myocardial infarction. *N Engl J Med* 1992;327:248–54.
  26. Munro J, Booth A, Nicholl J. Routine preoperative testing: a systematic review of the evidence. *Health Technol Assess* 1997;1(12):i–iv, 1–62.
  27. Donaldson LJ, Gray JAM. Clinical governance: a quality duty for health organisations. *Qual Health Care* 1998;7:S37–44.
  28. Murray CJL, Lopez AD. Evidence-based health policy—lessons from the Global Burden of Disease study. *Science* 1996;274:740–3.
  29. Boland M. What do people expect from their doctors? *World Health Forum* 1995;16:221–7.
  30. Thomas P. Competence, revalidation and continuing professional development—a business package. *Ann R Coll Surg Engl* 1999;81:13–5.
  31. Woolf SH, Grol R, Hutchinson A, Eccles M, Grimshaw J. Potential benefits, limitations, and harms of clinical guidelines. *Br Med J* 1999;318:527–30.
  32. Kricka LJ, Parsons D, Coolen RB. Healthcare in the United States and the practice of laboratory medicine. *Clin Chim Acta* 1997;267:5–32.
  33. Genç M, Mårdh P-A. A cost-effectiveness analysis of screening and treatment for *Chlamydia trachomatis* infection in asymptomatic women. *Ann Intern Med* 1996;124:1–7.
  34. Williams O. What is clinical audit? *Ann R Coll Surg Engl* 1996;78:406–11.
  35. Warren KS, Mosteller F, eds. Doing more harm than good: the evaluation of health care interventions. *Ann N Y Acad Sci* 1993;703:1–341.
  36. Bowling A, ed. Research methods in health. Investigating health and health services. Buckingham, UK: Open University Press, 1997:1–431.
  37. Nierenberg AA, Feinstein AR. How to evaluate a diagnostic marker test: lessons from the rise and fall of dexamethasone suppression test. *JAMA* 1988;259:1699–702.
  38. Irwig L, Tosteson ANA, Gatsonis C, Lau J, Colditz G, Chalmers TC, et al. Guidelines for meta-analyses evaluating diagnostic tests. *Ann Intern Med* 1994;120:667–76.
  39. Ohlsson A. Systematic reviews—theory and practice. *Scand J Clin Lab Invest* 1994;54:25–32.
  40. Greenhalgh T. How to read a paper: papers that report diagnostic or screening tests. *Br Med J* 1997;315:540–3.
  41. Silverstein MD, Boland BJ. Conceptual framework for evaluating laboratory tests: case finding in ambulatory patients. *Clin Chem* 1994;40:1621–7.
  42. Rink E, Hilton S, Szczepura A, Fletcher A, Fletcher J, Sibbald B, et al. Impact of introducing near-patient testing for standard investigations in general practice. *Br Med J* 1993;307:775–8.
  43. Storrow AB, Gibler WB. The role of cardiac markers in the emergency department. *Clin Chim Acta* 1999;284:187–96.
  44. Akiyama T, Pawitan Y, Greenberg H, Kuo C, Reynolds-Haertle R. Increased risk of death and cardiac arrest from encainide and flecainide in patients after non-Q-wave acute myocardial infarction in the Cardiac Arrhythmia Suppression Trial. *Am J Cardiol* 1991;68:1551–5.
  45. Price CP, Finney H. Developments in the assessment of glomerular filtration rate. *Clin Chim Acta* 2000;297:55–66.
  46. Guder WG. Preanalytical factors and their influence on analytical quality specifications. *Scan J Clin Lab Invest* 1999;59:545–50.
  47. Batstone G. Practising by the evidence: the role of pathology. *J Clin Pathol* 1997;50:447–8.
  48. Hobbs FDR, Delaney BC, Fitzmaurice DA, Wilson S, Hyde CJ, Thorpe GH, et al. A review of near patient testing in primary care. *Health Technol Assess* 1997;1(5):i–iv, 1–229.
  49. Rainey PM. Outcomes assessment for point-of-care testing. *Clin Chem* 1998;44:1595–6.
  50. Keffer J. Health economics of point-of-care testing. In: Price CP, Hicks JM, eds. Point-of-care testing. Washington: AACC Press, 1999:233–48.
  51. Vamvakas EC. Meta-analyses of studies of the diagnostic accuracy of laboratory tests. *Arch Pathol Lab Med* 1998;122:675–86.
  52. Moses LE, Shapiro D. Combining independent studies of a diagnostic test into a summary ROC curve: data-analytic approaches and some additional considerations. *Stat Med* 1993;12:1293–316.
  53. Guyatt GH, Oxman AD, Ali M, Willan A, McIlroy W, Patterson C. Laboratory diagnosis of iron-deficiency anemia: an overview. *J Gen Intern Med* 1992;7:145–53.
  54. Olatidoye AG, Wu AHB, Feng Y-J, Waters D. Prognostic role of troponin T versus troponin I in unstable angina pectoris for cardiac events with meta-analysis comparing published studies. *Am J Cardiol* 1998;81:1405–10.
  55. Lijmer JG, Mol BW, Heisterkamp S, Bossel GJ, Prins MH, van der Meulen JHP, et al. Empirical evidence of design-related bias in studies of diagnostic tests. *JAMA* 1999;282:1061–6.
  56. Mulrow CD, Linn WD, Gaul MK, Pugh JA. Assessing quality of a diagnostic test evaluation. *J Gen Intern Med* 1989;4:288–95.
  57. Kraemer HC, ed. Evaluating medical tests: objective and quantitative guidelines. Newbury Park, CA: SAGE Publications, 1992: 103–13.
  58. Jaeschke R, Guyatt G, Sackett DL. Users' guides to the medical literature. III. How to use an article about a diagnostic tests. A. Are the results of the study valid? Evidence-Based Medicine Working Group. *JAMA* 1994;271:389–91.
  59. van der Schouw YT, Verbeek AL, Ruijs SH. Guidelines for the assessment of new diagnostic tests. *Investig Radiol* 1995;30:334–40.
  60. Panzer RJ, Suchman AL, Griner PF. Workup bias in prediction research. *Med Decis Making* 1987;7:115–9.
  61. Valenstein PN. Evaluating diagnostic tests with imperfect standards. *Am J Clin Pathol* 1990;93:252–8.
  62. Begg CB. Biases in the assessment of diagnostic tests. *Stat Med* 1987;6:411–23.
  63. Rowley JM, Hampton JR. Diagnostic criteria for myocardial infarction. *Br J Hosp Med* 1981;26:253–8.
  64. Moore RA, Fingerova H. Evidence-based laboratory medicine: using current best evidence to treat individual patients. In: Price CP, Hicks JM, eds. Point-of-care testing. Washington: AACC Press, 1999:265–88.
  65. The Diabetes Control and Complications Trial Research Group. The effect of intensive treatment of diabetes on the development

- and progression of long-term complications of insulin-dependent diabetes mellitus. *N Engl J Med* 1993;329:977–86.
- 66.** Kendall JM, Reeves B, Clancy MJ. Point of care testing: randomized controlled trial of clinical outcome. *Br Med J* 1998;316:1052–7.
- 67.** Easterbrook PJ, Berline JA, Gopalan R, Matthews DR. Publication bias in clinical research. *Lancet* 1991;337:867–72.
- 68.** Dickersin K, Min Y-I, Meinert CL. Factors influencing publication of research results. Follow-up of applications submitted to two institutional review boards. *JAMA* 1992;267:374–8.
- 69.** Chalmers TC, Frank CS, Reitman D. Minimizing the three stages of publication bias. *JAMA* 1990;46:197–207.
- 70.** Haines A, Donald A. Getting research findings into practice. London: BMJ Publishing Group, 1998:1–160.
- 71.** Rawlins M. In pursuit of quality: the National Institute for Clinical Excellence. *Lancet* 1999;353:1079–82.
- 72.** Lord J, Littlejohns P. Evaluating healthcare policies: the case of clinical audit. *Br Med J* 1997;315:668–71.
- 73.** Sackett DL, Richardson WS, Rosenberg W, Haynes RB, eds. Evidence-based medicine. How to practice and teach EBM. London: Churchill Livingstone, 1997:1–250.
- 74.** Ramsey PG, Carline JD, Inui TS, Larson EB, Logerfo JP, Norcini JJ, et al. Changes over time in the knowledge base of practicing internists. *JAMA* 1991;266:1103–7.
- 75.** Davis DA, Thomson MA, Oxman AD, Haynes RB. Evidence for the effectiveness of CME. A review of 50 randomized controlled trials. *JAMA* 1992;268:1111–7.
- 76.** Muir Gray JA. Appraising the quality of research. In: Muir Gray JA, ed. Evidence-based healthcare. How to make health policy and management decisions. London: Churchill Livingstone, 1997:69–102.
- 77.** Price CP, Hicks JM. Point-of-care testing: An overview. In: Price CP, Hicks JM, eds. Point-of-care testing. Washington: AACC Press, 1999:3–16.
- 78.** Burnett D, ed. Understanding accreditation in laboratory medicine. London: ACB Venture Publication, 1996:1–311.
- 79.** Klein R. Competence, professional self regulation, and the public interest. *Br Med J* 1998;316:1740–2.