



Current Evidence and Future Perspectives on the Effective Practice of Patient-Centered Laboratory Medicine

Mike J. Hallworth,^{1*} Paul L. Epner,² Christoph Ebert,³ Corinne R. Fantz,⁴ Sherry A. Faye,⁵ Trefor N. Higgins,⁶ Eric S. Kilpatrick,⁷ Wenzhe Li,⁸ S.V. Rana,⁹ and Florent Vanstapel,¹⁰ on behalf of the IFCC Task Force on the Impact of Laboratory Medicine on Clinical Management and Outcomes

BACKGROUND: Systematic evidence of the contribution made by laboratory medicine to patient outcomes and the overall process of healthcare is difficult to find. An understanding of the value of laboratory medicine, how it can be determined, and the various factors that influence it is vital to ensuring that the service is provided and used optimally.

CONTENT: This review summarizes existing evidence supporting the impact of laboratory medicine in healthcare and indicates the gaps in our understanding. It also identifies deficiencies in current utilization, suggests potential solutions, and offers a vision of a future in which laboratory medicine is used optimally to support patient care.

SUMMARY: To maximize the value of laboratory medicine, work is required in 5 areas: (a) improved utilization of existing and new tests; (b) definition of new roles for laboratory professionals that are focused on optimizing patient outcomes by adding value at all points of the diagnostic brain-to-brain cycle; (c) development of standardized protocols for prospective patient-centered studies of biomarker clinical effectiveness or extraanalytical process effectiveness; (d) benchmarking of existing and new tests in specified situations with commonly accepted measures of effectiveness; (e) agreed definition and validation of effectiveness measures and use of checklists for articles submitted for publication. Progress in these areas is essential if we are to demonstrate and enhance the value of laboratory medicine and prevent valuable information being lost in meaningless data. This requires effective

collaboration with clinicians, and a determination to accept patient outcome and patient experience as the primary measure of laboratory effectiveness.

© 2014 American Association for Clinical Chemistry

Clinical laboratories provide valuable services to aid patient diagnosis and management, but systematic evidence of laboratory medicine's specific contribution to the overall process of healthcare is difficult to find. An understanding of the value of laboratory medicine and the various factors that influence it is of paramount importance to ensuring that the service is provided and used optimally to improve patient care and that resources (technological, financial, and human) are not wasted, inappropriately deployed, or unwisely constrained.

Many articles seeking to promote the value of laboratory medicine have made use of what has become known as the “70% claim.” This is presented in various forms, most commonly that “Laboratory medicine influences 70% of clinical decisions” or minor variations around this figure. Regrettably, this estimate was based on unpublished studies and anecdotal observations and cannot now be objectively verified (*1*). Furthermore, a single “headline number,” however accurate, is unhelpful for a detailed understanding of the contribution of laboratory medicine to quality healthcare.

More specific and evidence-based indices of the incremental value of laboratory medicine are needed, which require better understanding of the mechanisms by which value is added (or reduced) along with well-designed clinical studies. The IFCC Task Force on the Impact of Laboratory Medicine on Clinical Management and Outcomes was established in 2012 to evaluate the available evidence supporting the impact of laboratory medicine in healthcare, and to develop the study design for new retrospective and prospective studies capable of generating evidence of the contribution made by laboratory medicine. This review will summarize existing evidence and indicate the gaps in our understanding. It will also identify deficiencies in current utilization, suggest potential solutions, and offer a vision of a future state

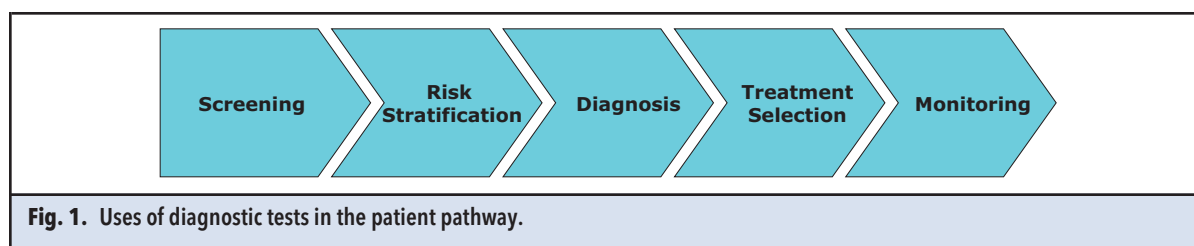
¹ Department of Clinical Biochemistry, Royal Shrewsbury Hospital, Shrewsbury, Shropshire, UK; ² Paul Epner LLC, Evanston, IL; ³ Roche Diagnostics GmbH, Penzberg, Germany; ⁴ Labsource LLC, Atlanta, GA; ⁵ Beckman Coulter Eurocenter SA, Nyon, Switzerland; ⁶ DynaLIFE Dx, Edmonton, Alberta, Canada; ⁷ Department of Clinical Biochemistry, Hull Royal Infirmary, Hull, UK; ⁸ NMS Labs, Willow Grove, PA; ⁹ Department of Gastroenterology, PGIMER, Chandigarh, India; ¹⁰ University Hospitals Leuven-Gasthuisberg, Leuven, Belgium

* Address correspondence to this author at: 17 Fairlawn Ave., Meole Brace, Shrewsbury SY3 9QQ, UK. E-mail: mikejhallworth@gmail.com.

Received September 11, 2014; accepted December 5, 2014.

Previously published online at DOI: 10.1373/clinchem.2014.232629

© 2014 American Association for Clinical Chemistry



in which laboratory medicine is used optimally to support patient care, enhance patient safety, and improve outcomes.

Definition and Purpose of Laboratory Medicine

Laboratory medicine may be defined (*Mosby's Medical Dictionary*) as “the branch of medicine in which samples of tissue, fluid, or other body substance are examined outside of the person,” either in a laboratory or at the point of care (near-patient testing). It encompasses the traditional disciplines of clinical chemistry, toxicology, hematology, immunology, microbiology (including serology and virology), anatomical pathology, cytology, molecular pathology, and cytogenetics. The global in vitro diagnostics market was valued at \$49.2 billion in 2012, projected to grow at a rate of 7% from 2012–2017 (2). This probably corresponds to 10–15 billion tests per year worldwide (authors' estimate).

Laboratory medicine supports the interaction between patients and physicians by providing relevant data, increasing the likelihood of making the most appropriate decisions for the optimum care of the individual. Gambino (3) and Lundberg (4) defined the interaction in terms of the total testing process (TTP)¹¹ or “brain-to-brain loop.”

The Value of Laboratory Medicine

Isolating the value of the testing process from the overall process of patient management is complicated. Value can be defined clinically or economically. Clinical value is linked to the improvement of health-related outcomes, whereas economic value is tied to cost-efficiency or -effectiveness (5). The former is not specific to individual countries or healthcare systems, except insofar as it depends on the spectrum of patients presenting to those systems. The latter will depend on the specific setting in

which the testing process is to be deployed and the range and cost of diagnostic tools available. Economic analyses of value are thus difficult to transfer between settings. We will focus on clinical value in its broadest sense, because there is no such thing as cost-effectiveness for a clinically ineffective process.

We propose an approach to measuring value in which the net value of a testing process is defined as delivered benefits minus delivered harm (undesirable effects of testing). In this model, net value is maximized by increasing the benefits or reducing harm.

The benefits derive from the provision of objective data about patient health that enable screening of populations for detection of asymptomatic disease, stratification of risk, specific diagnosis of a patient's condition, selection and monitoring of appropriate therapy, prediction and early detection of adverse treatment outcomes, and assessment of prognosis (Fig. 1).

Existing Evidence of Value

It is outside the scope of this paper to offer a comprehensive review of all studies which address the value of laboratory medicine. Instead, we present examples that demonstrate benefit (Table 1). Further evidence can be found in a US national status report (13).

POINT-OF-CARE TESTING

Point-of-care testing (POCT) is a specialized situation with unique attributes. “The close proximity of the patient, a clinical provider with a question, and the technology to rapidly answer that question” leads to quicker decisions and actions compared to centralized testing (14). This has been demonstrated with varied conditions, symptoms, and settings (15–18). These unique benefits are accompanied by unique concerns which are discussed later.

COST-EFFECTIVENESS

Although cost-effectiveness is not a primary focus of this review, a brief discussion of this topic is relevant. The cost-effectiveness of laboratory testing has been studied in various settings [reviewed by Hernandez et al. (19) and Fang et al. (20)].

¹¹ Nonstandard abbreviations: TTP, total testing process; POCT, point-of-care testing; ICER, incremental cost-effectiveness ratio; QALY, quality-adjusted life year; Hb A_{1c}, hemoglobin A_{1c}; ICHCLR, International Consortium for Harmonization of Clinical Laboratory Results; RCT, randomized controlled trial; PSA, prostate-specific antigen.

Table 1. Examples of existing evidence of the value of laboratory testing.

Purpose of testing	Value of testing	Examples
Screening	Prenatal and neonatal screening has well-proven clinical value and cost-effectiveness compared to the cost of caring for an individual who may require lifelong support [Lewin Group (6)].	Down syndrome, hypothyroidism, phenylketonuria, medium-chain acyl-CoA dehydrogenase deficiency.
	Screening for cancer in appropriately targeted populations has proven effectiveness in reducing morbidity and mortality.	Fecal occult blood in bowel cancer [Mandel et al. (7)], PAP smears.
Risk stratification	Laboratory testing is integral to risk assessment of future complications and the need for therapeutic intervention.	Cardiovascular disease (lipids, troponins), heart failure (B-type natriuretic peptide), diabetes (Hb A _{1c}), and bone disease (hydroxylated vitamin D).
Diagnosis	Laboratory medicine is central to the diagnostic process for a wide range of conditions. In the emergency room setting, laboratory testing is requested in more than 41% of all visits [CDC (8)].	Evaluation of unexplained coma, chest pain, abdominal pain, shortness of breath.
	Many diseases are defined by the results of laboratory tests or cannot be definitively diagnosed without them.	Diabetes mellitus, lipid disorders, all cancers, most endocrine disorders, infectious diseases, thrombotic disorders, and the majority of inherited diseases.
	Laboratory tests provide early warning of developing disease, allowing rapid intervention.	Creatinine in acute kidney injury [Kellum and Lameire (9)], C-reactive protein and procalcitonin in infection/sepsis.
	The growing significance of laboratory testing may also reflect physicians' increasing reliance on objective data.	
	Laboratory medicine is being used to subtype disease populations, with important implications for treatment. Many specific cancer therapies have "companion diagnostics." These tests provide information on an individual's suitability to receive a particular treatment or dose, and the likelihood of response.	Serum creatinine or estimated glomerular filtration rate for renally excreted drugs, red cell thiopurine methyltransferase in azathioprine therapy, KRAS (Kirsten rat sarcoma viral oncogene homolog) gene mutation detection for anti-epidermal growth factor receptor therapy.
Treatment selection:	Effective antibiotic therapy is heavily dependent on culture and sensitivity testing.	
	More widely, therapeutics is being transformed by pharmacogenetics, allowing improved selection of effective drugs and their dosage, and avoidance of predictable drug interactions [Urban and Goldstein (10)], an excellent example of the clinical and economic value of laboratory testing.	
Monitoring of treatment response	Ongoing measurement of a test or tests with a view to evaluating a particular intervention, enabling continuous optimization of management depending on the results.	Hb A _{1c} in diabetes, thyroid-stimulating hormone in hypothyroidism, measurement of drug concentrations such as antibiotics, anticonvulsants, and immunosuppressives, international normalized ratio testing, evaluation of HIV viral load.

Continued on page 592

Fang et al. (20) reviewed the Tufts Medical Center Cost-Effectiveness Analysis registry for studies involving laboratory diagnostics. A total of 141 cost-utility studies up to 2008 were included, containing 433 separate incremental cost-effectiveness ratios (ICERs). The ICER is

defined by: $(C1 - C2)/(E1 - E2)$, for which C1 and E1 are the cost and effect in the intervention or treatment group and C2 and E2 are the cost and effect in the control group. Information about test accuracy was clearly reported for 63% of the studies, but test safety or risk of

Table 1. Examples of existing evidence of the value of laboratory testing. (Continued from page 591)

Purpose of testing	Value of testing	Examples
	Laboratory data underpins the practice of evidence-based medicine and the development of clinical practice guidelines for the management of many conditions; 23% of the quality indicators assessed in the RAND Corporation's analysis of the quality of care delivered to US adults involve laboratory tests [McGlynn et al. (11)]. Out of 1230 evidence-based clinical practice guidelines for the 23 main condition/disease guidelines defined by the National Guidelines Clearinghouse, 37% focused on or involved laboratory tests [Lewin Group (12)].	Cardiovascular disease: All 2012 guidelines for the European Society of Cardiology (ESC) were examined by the present authors to determine the role of laboratory testing. Of 34 cardiovascular clinical guidelines published on the ESC website (http://www.escardio.org/guidelines-surveys/esc-guidelines/Pages/GuidelinesList.aspx) in December 2012, laboratory tests were integrally related to diagnosing disease in 17, and to monitoring and follow-up in 25. In 14 guidelines, laboratory tests were required for both diagnosis and management. Only 5 guidelines made no specific recommendations about involvement of laboratory testing in the decision-making process.

testing were mentioned for only 10%, and the potential value or harm of testing unrelated to treatment options was considered in only 7% of studies. Fourteen percent of the individual ICERs were cost-saving gains at less cost—the classic “win-win” situation. Gains are expressed in terms of quality-adjusted life years (QALYs), for which years of life gained by an intervention are multiplied by a utility value associated with the quality of a given state of health (1 year in perfect health = 1 QALY). A further 43% had cost-utility estimates below the commonly accepted threshold of \$50 000 per QALY gained. The review made recommendations for standardization of cost-utility methodology and concluded that the literature reveals many areas in which testing represents good value for money in terms of QALYs gained. However, it did not consider the costs of delayed or wrong diagnosis as a result of not performing the appropriate test—information that would have produced even more positive findings. Most importantly, it is these downstream impacts that swamp the benefits of reduced laboratory expenses because all of laboratory testing is typically only 3%–5% of national healthcare costs (21).

Potential Mechanisms for Harm

Despite the many benefits, it must also be recognized that inappropriate use of laboratory testing or test results can result in poorer patient care. Sources of harm include inappropriate medical interventions driven by misdiagnosis, unnecessary treatment, or additional diagnostic procedures following detection of a “disease” which turns out to be clinically insignificant, or failure to receive appropriate treatment associated with a correct diagnosis (22, 23).

Astion et al. (24) were among the first to recognize that errors in the testing process can harm patients and reduce the effectiveness of laboratory medicine. They developed a classification system to prioritize quality improvement initiatives and improve effectiveness. More recently, Epner et al. (23) have described 5 ways in which the testing process can lead to diagnostic error, all of which can impact an individual patient's care and reduce the real-world effectiveness of testing programs (Table 2). If overall patient harm, rather than diagnostic error alone, is considered, a sixth cause could be added, to cover situations in which the testing process itself causes harm to patients (14).

Although effort is still needed in the area of inaccurate results (cause 5), it is now by far the least frequent cause of error (25). Reducing diagnostic errors associated with causes 1–4 should now be the primary focus of a

Table 2. Classification of causes of diagnostic error related to testing.^a

1. An inappropriate test is ordered.
2. An appropriate test is not ordered.
3. An appropriate test result is misinterpreted or misapplied.
4. An appropriate test is ordered, but is delayed or misrouted within the testing process and is not available where it is needed at the time it is needed.
5. The result of an appropriately ordered test is inaccurate.

^a Reproduced from Epner et al. (23).

patient-centered approach to improving laboratory testing (26) and will be considered in detail below.

Earlier we mentioned the specialized case of POCT and described the benefit of faster decision-making. However, the 5 causes of potential harm are also present in this situation but manifest themselves differently and with different frequency and potential for harm. Unfavorable technical performance, lower levels of training of testing personnel, different QC protocols, and less time and opportunity for error detection associated with the faster decisions all demand a new approach to considering the value of POCT (14).

Current State of Laboratory Medicine

The emphasis of laboratory medicine research over the last 30 years or so has been focused on improving performance within the laboratory. In response to the obvious harm caused by poor agreement of results between laboratories, much progress has been made on the standardization and harmonization of methods and results [e.g., lipids, hemoglobin A_{1c} (Hb A_{1c}), testosterone, creatinine] (27). The International Consortium for Harmonization of Clinical Laboratory Results (ICHCLR) (www.harmonization.net) has been formed to prioritize and organize global activities to achieve harmonization of test results, and there have been initiatives on the standardization of reference intervals to ensure consistency of interpretation and on process optimization within the laboratory. This work is in part driven by the data from proficiency testing or external quality assurance. Improvements in internal QC have helped improve the efficiency of laboratory internal processes, and platform consolidation and more extensive automation have improved cost efficiency. In addition, laboratories have devoted considerable efforts to improving the physician–laboratory interface, with computerized physician order entry systems that, in some cases, provide context-specific help on appropriate test requesting. There has also been more emphasis on consistency in result reporting, including when and how critical results are notified to the requesting clinician (28). These changes have been driven by the current healthcare macroenvironment, specifically the need to control healthcare costs, enhance patient safety, and deliver greater transparency.

Outcome Studies in Laboratory Medicine

Although all this work has been extremely valuable and has improved laboratory efficiency, it has not often addressed the question of clinical effectiveness—the question of whether testing improves patient outcome. Outcomes have been defined as results of medical interventions in terms of health or cost (29). The primary outcome measures of any healthcare system were set out in the Institute of

Medicine's 2001 report “Crossing the Quality Chasm” (30)—STEEEP (safety, timeliness, effectiveness, equity, efficiency and patient-centeredness).

Outcome studies must be distinguished from studies of the clinical validity or prognostic accuracy of tests. Prognostic accuracy studies ask the question “is the result of the test associated with an outcome of interest?” whereas outcome studies ask whether the use of the test is associated with improved outcomes—do patients who have the test fare better than patients who do not? The mere fact of having the test result available does not make any contribution to the health or wellbeing of the individual patient. For the test to confer benefit, someone (clinician, other health professional, or patient) must do something useful or effective on the basis of the result. The value of laboratory medicine must always be considered in the context of defined care pathways. A new test or biomarker may show diagnostic value when considered by itself, but fail to add significant value when considered in the context of the overall diagnostic workup for the patient. These situations can occur when different tests measure the same underlying pathological process, thus providing similar diagnostic information (31) (e.g., heart-type fatty-acid binding protein in early detection of myocardial infarction).

Improved outcomes may be seen in terms of faster or more accurate diagnosis, improved treatment selection, avoidance of misdiagnosis or adverse treatment consequence, improved patient flow, and improved patient satisfaction or quality of life (32). Although studies normally concentrate on the effects of testing strategies on clinical decision-making and downstream management, Bossuyt and McCaffrey (33) have provided a useful summary of other patient outcomes such as cognitive, emotional, social, and behavioral effects, and Lee et al. (34) have shown how the value of a diagnostic test can be analyzed as medical, planning (the effect on patients' ability to make better decisions), and psychic (how the information from the test affects the patient's sense of self).

True outcome studies are difficult to do. This is because the measured outcomes may be many steps beyond the performance of the test and affected by many other intermediate variables. For example, the clinician may be inconsistent in his/her use of the decision threshold, or there may be no effective therapy for the condition. The success of outcomes studies typically requires standardized and agreed responses to the test results, to preserve the link with the desired outcome. Even then, the effectiveness of the intervention initiated by a particular test result may vary across patient populations, reducing the effectiveness of the linkage. Other problems that have affected previous population-based outcomes studies of laboratory testing have included lack of standardization of data collection and reporting methods,

lack of agreement on appropriate data analysis, small sample sizes, missing or incomplete data in patients' medical records, and the high cost of collecting outcomes data.

What Needs to Be Done?

We need studies which can form a proper assessment of the impact of laboratory testing on patient outcome and will lead to improved outcomes with optimum utilization of the personnel, processes, and technology involved. The 2009 Lewin Group report on the value of laboratory tests (7) notes that, to improve outcomes, a laboratory test must be appropriately ordered and conducted and the results must be returned and retrieved on a timely basis and correctly interpreted and must affect a decision for further diagnosis and treatment that directly improves outcome.

There are 2 fundamental issues here. To clearly demonstrate the link between a particular testing strategy and a defined outcome, first the test needs to be used appropriately—requested on the right patient population at the right stage in the process, analyzed promptly, the result returned to the right place, and the appropriate action taken within the optimum time frame. Second, the study design used to measure the link between testing and outcome needs to be rigorously defined and properly implemented. In addition, allowance must be made for the Hawthorne effect—good practice may not persist when not enforced by the study context, and subsequent audit is essential.

ENSURING APPROPRIATE UTILIZATION

We consider the ensuring of appropriate test utilization in terms of the 5-cause model described above (Table 2).

Causes 1–3 relate to appropriate ordering and interpretation of testing. The improvement of test ordering patterns must be directed to better patient care and designed to produce a more effective diagnostic process. Test utilization initiatives that are driven solely by financial rather than patient care considerations usually are either ineffective or produce short-lived financial gains. Obviously, payers do not want to fund irrelevant, wasteful, or redundant testing, but they must recognize that testing plays a key role in supporting rapid and effective diagnosis and treatment. The focus must be on cost per specific outcome, rather than cost per test.

Hickner et al. have recently described the challenges faced by primary care physicians in ordering and interpreting laboratory tests (35). Appropriate ordering depends on effective clinician education and support, and the laboratory has a key role in designing physician–laboratory interface procedures [whether they be CPOE (computerized physician order entry) or manual request forms] that make it easy to order the right tests in a

defined situation and more difficult to request inappropriate ones (36). Solutions may include devising standard protocols for specific presentations, and problem-based ordering, in which the clinician indicates the problem or question to be answered and the laboratory defines the test strategy on the basis of clinical findings and the results of first-line tests and provides interpretation of the results (37, 38). Appropriate use of laboratory testing is critically dependent on the pretest probability of disease in a given situation—pretest disease probabilities that are very high or very low are both likely to result in poor utilization of testing.

Tackling underutilization (cause 2) is as important as reducing overutilization in improving outcomes. A recent 15-year metaanalysis of inappropriate laboratory testing concluded that underutilization is widespread but understudied (39). In this analysis, the overall mean rate of overutilization of laboratory testing across 42 studies was 20.6% (95% CI, 16.2%–24.9%), whereas the overall mean rate of underutilization of testing was more than twice as high (44.8%; 95% CI, 33.8%–55.8%). An emphasis on ordering the “right test” for a particular clinical presentation leads to confirmation of “obvious” diagnoses. But in primary diagnosis the important issue is not to confirm the obvious, but rather not to miss the unexpected (40). An analysis of US malpractice claims (41) showed that in 55% of malpractice claims relating to a missed diagnosis, failure to order the correct diagnostic test was a factor in the outcome.

Misinterpretation or misapplication of results (cause 3) can also result from a failure to understand the limitations of the test (sensitivity, specificity, effect of interfering substances) or the lack of essential information needed to interpret the result in the appropriate clinical context. As Black described in 1998 in the context of radiology, the ability to detect smaller abnormalities axiomatically tends to increase the prevalence of any given disease (42). More sensitive diagnostic tests clearly allow the detection of less severe forms of diseases or disorders, which is not in itself a bad thing but does require serious thought about when to use diagnostic labels and therapeutic approaches traditionally deployed against more serious presentations of disease (23, 43). For example, the prostate-specific antigen (PSA) test may result in diagnosis of indolent prostate cancer which will have few or no clinical consequences during the patient's lifetime. The uncertainty about the aggressive nature of individual prostate cancers has led to overtreatment and considerable harm when providers choose treatment interventions for all patients with a positive PSA test. This is not an argument against the value of laboratory testing per se, but emphasizes the importance of appropriate utilization.

Laboratory professionals naturally wish to provide the most accurate and sensitive tests that the technology

allows, but they must also accept the obligation to partner with clinicians to determine the relevance of the results obtained with regard to diagnostic, treatment, or monitoring decisions. This involves meaningful communication of factors such as predictive value and the significance of changes in results. The benefits and associated problems of increasing the sensitivity of troponin assays are a case in point (44).

Appropriate interpretation can be aided by quality interventions such as the cholesterol standardization program in the US (45). By targeting treatment more effectively, improvement in lipid standardization has contributed to the marked reduction in heart disease deaths seen since 1980. Using very conservative assumptions (that only 0.5% of the cholesterol-related improvement in mortality is attributable to the standardization program and assigning a value of \$50 000 per life-year saved), the authors estimated the total cost benefit of lipid standardization at \$338 million, vastly greater than the program cost (\$2.7 million in 2007). With less-conservative assumptions, the cost benefit runs into billions of dollars, although the cause-and-effect relationship cannot be unequivocally established.

The World Alliance for Patient Safety (46) has identified poor test follow-up (cause 4) as one of the major processes contributing to unsafe patient care, and the safety implications of missed test results for hospitalized patients have been reviewed by Callen et al. (47). Although the number of research studies is limited and variations in study methodology prevented robust comparisons, there is strong evidence that the problems caused by missed test results are considerable, including missed cancer diagnoses and missed *Chlamydia* infection leading to pelvic inflammatory disease and avoidable patient deaths (47). In 1 typical study, 29% of urgent biochemistry results requested over a 6-month period were never accessed electronically (48).

Appropriate communication of critical results to the right person at the right time—and avoidance of “alert fatigue” caused by inappropriate communication of abnormal but not critical results—has also been shown to have an high impact on patient outcome (49). Kost and Hale (28) have reviewed global practice in notification of critical results and recommended improved standardization of reporting in this area.

The problem of inaccurate results (cause 5) has received much attention within laboratories over the last 40 years, and in general the analytical specificities and limitations of tests are well understood—at least by laboratory professionals, though not always by the users of test results. Action is being taken to rectify the remaining issues through the ICHCLR, as mentioned above. However, insufficient attention has been paid to other aspects of the TTP, notably preanalytical factors which may have a profound effect on the accuracy of test results. Com-

mon examples are the incidence of hemolysis (in samples from outlying locations or due to collection practices in the emergency department) and contamination of blood cultures. The effects of these factors can have major implications for patients, including inappropriate admission to the hospital (50) or delayed care due to the need to repeat the test. Improving the efficiency of pre- and postanalytical processes is part of the discipline of laboratory medicine and cannot be left to clinical professionals. It requires close collaboration between laboratory professionals and service users.

DESIGNING BETTER EVALUATIONS

As stated above, the evidence for the clinical utility of diagnostic tests in the peer-reviewed literature has been poor, with the majority of reports dealing with the analytical or technical performance of tests and their association with disease states. There has been little emphasis on validated measures which link testing with patient or population outcomes. Assessment of laboratory tests has focused on reporting diagnostic accuracy in terms of sensitivity and specificity with respect to the population studied (51) and making extensive use of potentially misleading techniques such as ROC curves. Stakeholders (such as care providers, payers, and policy makers) have, perhaps unsurprisingly, been resistant to accepting the value of these studies, but generally respond favorably to studies providing hard evidence of benefit in overall patient outcome (52, 53).

The weaknesses of dichotomized sensitivity and specificity studies are well understood. They are often heavily biased by the effects of comparison with non-gold standard classification methods and by the compounding effects of prevalence and disease spectrum (54). Comparison of diseased patients with matched healthy controls may help to prove statistical significance but does not reflect the normal clinical scenario in which the distinction is not between healthy and diseased, but between patients with a specific disease and others with a similar presentation but a different disease.

It is only lately that the principles of evidence-based medicine have been applied—not always effectively or consistently (55). New tests such as cardiac troponins, natriuretic peptides, procalcitonin, and fecal calprotectin have been introduced using robust evidence of improved outcomes, but much current “routine” testing has not undergone the same scrutiny. There are 2 missed opportunities here—increasing the value of older tests as a result of better understanding of how to use them effectively (e.g., serum creatinine in the prediction of acute kidney injury) and eliminating outmoded older tests because of new evidence that they are useless at best or harmful at worst.

The most difficult part of producing evidence is to pose the right research question (56). The question

must be formulated in a way that highlights the variables that will impact upon the effectiveness of the investigation in the clinical scenario being considered, using the PICO framework (Population, Intervention, Comparator, Outcome). Several variants have been described (56).

Once the link between intervention and outcome has been rigorously defined, various evaluation methods can be employed [see the recommendations of the European Federation of Clinical Chemistry and Laboratory Medicine's Test Evaluation Working Group (57)]. Bossuyt et al. (58) describe how comparative randomized controlled trials (RCTs) can be used in such scenarios to estimate clinical utility and also to show how diagnostic accuracy can be linked to clinical utility. From the same group, Moons et al. (31) have provided a useful overview of methods to quantify the added value of a new diagnostic test. Mallett et al. (59) also highlight different ways in which results from diagnostic test accuracy studies can be presented and interpreted.

The usual emphasis on RCTs in clinical research is a challenge to the demonstration of the value of diagnostic testing, given the confounding effects of intervening decisions, the variable effectiveness of downstream management and environmental factors between testing and ultimate outcome, and the costs and time (years or even decades) needed to assess the impact of outcomes using RCT. For these reasons, the Lewin Group's 2010 report *Laboratory Medicine and Comparative Effectiveness Research* (60) stated that RCTs should not be the default design for comparative effectiveness research in laboratory medicine. However, Lord et al. (61) recommend using the principles of RCT design in the design of test evaluations, to identify the types of evidence that will be required for effective evaluation. Trikalinos et al. (62) have explored the use of decision-analytical simulation models, often used to inform reimbursement decisions for drugs and vaccines, as a tool for assessing the value of diagnostic tests.

The STARD (Standards for Reporting of Diagnostic Accuracy) initiative (63) provided a checklist designed to improve the quality of reporting of studies of diagnostic accuracy and the CONSORT (Consolidated Standards Of Reporting Trials) statement (64) is a useful guide to designing RCTs studying the effects of interventions on outcomes. Design requirements for studies assessing the utility of diagnostic tests have been described in detail in a recent AHRQ (Agency for Healthcare Research and Quality) publication (65), and the QUADAS-2 (the revised Quality Assessment of Diagnostic Studies) checklist (66) provides a generic tool for the quality assessment of primary studies within a diagnostic review. A very useful checklist to determine clinically important differences between test-treatment pathways of new and existing diagnostic tests has recently

Table 3. European Group on Tumour Markers 4-phase approach to evaluation of tumor markers.^a

Phase 1	Assessment of biomarker kinetics and correlation with tumor burden.
Phase 2	Assessment of the ability of the biomarker to identify, exclude and/or predict a change in disease status.
Phase 3	Assessment of the effectiveness of biomarker-guided intervention by measuring patient outcome in randomized trials.
Phase 4	Audit of the long-term effects after the biomarker has been introduced into the patient care pathway.

^a Adapted from Soletormos et al. (70).

been published by Ferrante di Ruffano et al. (67). The Institute of Medicine has also published standards for systematic reviews which address patient outcome (68). We recommend that such checklists should be more widely used by journals at the level of submission and at the level of the referee process and that the submitted checklists should become part of the submitted papers, or available as complementary material. This will improve the quality of research and of the refereeing process and will enhance the ability of metaanalyses to compile existing research.

The EQUATOR (Enhancing the Quality and Transparency of Health Research) group (69) has produced a catalogue of recommendations for promoting transparent and accurate reporting of medical research. In response to this, the European Group on Tumor Markers has published a proposal on evaluation of new tumor markers (70). This describes a 4-phase approach, similar to the process used by the US Food and Drug Administration and others for the evaluation of new drugs (Table 3).

We strongly recommend the extension of this approach into other areas of laboratory medicine. Adoption of the terminology (phase 1 studies = characterization of the biomarker; phase 2 studies = sensitivity, specificity, and predictive value in defined populations; phase 3 studies = trials on patient outcome; phase 4 studies = audit following implementation) would be immensely helpful when reporting biomarker studies. Such studies should be registered before initiation and published on completion, because failure to publish and selective reporting have been shown to be as prevalent in test accuracy studies as in other areas of science (71).

In addition to the outcomes value of specific tests in defined circumstances, we also need better evaluations of the value of the appropriate use of laboratory medicine

overall. The CDC-sponsored Laboratory Medicine Best Practices Initiative has described the methodology for identifying effective practices associated with improved healthcare quality outcomes (72), focusing on the pre- and postanalytical aspects of the TTP, where most errors occur. Such studies should include assessment of the effect on patient outcomes of consultative roles (38), the use of decision-support algorithms, the value of interpretative comments (73), and other practices to improve the overall efficiency of the TTP. Changes in laboratory practice, however well intentioned, also need to be investigated for potential effects on patient outcome. A good example is recent work by Kilpatrick et al. (74) showing that the change to millimole per mole units for reporting Hb A_{1c} measurements in the UK had no detrimental effect on patient outcome as measured by overall diabetic control.

The Vision—Patient-Centered Laboratory Medicine

Medicine in the 21st century needs a flexible information resource that facilitates selection of the right test on the right patients at the right time, with results delivered in a timely fashion to the right place accompanied by context-specific interpretation and, where appropriate, linked to guidance on agreed action to be taken.

To achieve this, work is required in 5 areas:

1. Improved utilization of existing and new tests. This requires determination of optimum testing strategies on the basis of patients' presenting complaints; development of interventions to support appropriate test ordering/requesting; proper sample collection, transport, and storage; effective strategies for transmission of test results; agreement on clinically appropriate triggers for critical result notification (28); and consultative services and comments to ensure that results are properly applied. Guidelines on testing strategies should be devised, rigorously tested, and published, with emphasis on how they can be effectively applied (75).
2. Defining new roles for laboratory professionals that are focused on optimizing patient outcomes by adding value at all points of the diagnostic brain-to-brain cycle and auditing the effectiveness of these roles and the overall diagnostic process.
3. Development of standardized protocols for prospective patient-centered studies of biomarker clinical effectiveness or extraanalytical process effectiveness (57, 58).
4. Benchmarking of existing and new tests in specified situations with commonly accepted measures of effectiveness using the models described above, including postimplementation audit (57, 70). This must include the effects of pre- and postanalytical components of the testing process and must consider overall impact of the testing process on all relevant clinical outcomes.

5. Agreed definition and validation of effectiveness measures and use of checklists for articles submitted for publication [see references (23) and (67)].

Call to Action

Substantial progress in the areas listed above is essential if we are to demonstrate and enhance the value of laboratory medicine and prevent the loss of valuable information in a miasma of meaningless data. Laboratory medicine professionals who choose to work in these areas will play a leading role in improving patient care and developing laboratory medicine as a clinical specialty, rather than a number-generating service. This requires effective collaboration with clinicians, and a determination to accept patient outcome and patient experience as the primary measure of laboratory effectiveness. Outcomes research must be properly funded by government or commissioning agencies, and the in vitro diagnostics industry has an important role to play in jointly facilitating and coordinating appropriate clinical studies. They cannot, however, deliver what is required without leadership from and partnership with the laboratory medicine community. The time is now right for laboratory medicine specialists in all countries to take up the challenge to look outside the laboratory and forge effective links with multidisciplinary teams that seek to optimize clinical outcomes and patient experiences in an efficient and cost-effective way (76).

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 author disclosure form. Disclosures and/or potential conflicts of interest:

Employment or Leadership: M.J. Hallworth, IFCC; P.L. Epner, Clinical Laboratory Management Association, and Society to Improve Diagnosis in Medicine; C. Ebert, Roche Diagnostics GmbH; C.R. Fantz, AACC; T.N. Higgins, DynaLIFEDx.

Consultant or Advisory Role: P.L. Epner, Silicon BioDevices, Integrated Project Management Company, Inc., Research Triangle International, Inc., Battelle, Glenwood Capital Management, Abbott Laboratories, and De Gruyter (*DIAGNOSIS*, Associate Editor); T.N. Higgins, Becton Dickinson.

Stock Ownership: P.L. Epner, Abbott Laboratories (stock options) and Silicon BioDevices.

Honoraria: P.L. Epner, Orchard Software, G2 Intelligence, and Saudi Arabia Quality Council; T.N. Higgins, Bio Rad.

Research Funding: None declared.

Expert Testimony: None declared.

Patents: None declared.

Other Remuneration: M.J. Hallworth, AACC, IFCC, ACB, and Cogara Conferences; P.L. Epner, ASCLS chapters, CLMA chapters,

Association of Pathology Informatics, Virginia Commonwealth University, CDC, IFCC, and Society to Improve Diagnosis in Medicine; C.R. Fantz, IFCC travel for committee meeting.

Acknowledgments: The authors are grateful to P. Gillery and members of the IFCC Scientific Division for helpful comments on the draft manuscript.

References

- Hallworth MJ. The 70% claim: what is the evidence base? *Ann Clin Biochem* 2011;48:487–8.
- Research and Markets. In vitro diagnostics (IVD) market, techniques and applications. Dublin: Research and Markets;2013. http://www.researchandmarkets.com/publication/pfdniq/in_vitro_diagnostic_ivd_market_technique (Accessed September 2014).
- Gambino SR. Met and unmet needs of the automated clinical laboratory. *Trans N Y Acad Sci* 1970;32(7 Series II):816–20.
- Lundberg GD. Acting on significant laboratory results. *JAMA* 1981;245:1762–3.
- Drummond MF, Sculpher MJ, Torrance GW, O'Brien BJ, Stoddart GL. *Methods for the economic evaluation of health care programmes*. 3rd ed. Oxford: Oxford Medical Publications; 2005.
- The Lewin Group. The value of laboratory screening and diagnostic tests for prevention and health care improvement. <http://www.lewin.com/publications/publication/403/> (Accessed September 2014).
- Mandel JS, Church TR, Ederer F, Bond JH. Colorectal cancer mortality: effectiveness of biennial screening for fecal occult blood. *J Natl Cancer Inst* 1999;91:434–7.
- National hospital ambulatory medical care survey: 2010 Emergency department summary tables. http://www.cdc.gov/nchs/data/ahcd/nhamcs_emergency/2010_ed_web_tables.pdf (Accessed September 2014).
- Kellum JA, Lameire N. Diagnosis, evaluation, and management of acute kidney injury: a KDIGO summary (Part 1). *Crit Care* 2013;17:204.
- Urban TJ, Goldstein DB. Pharmacogenetics at 50: genomic personalization comes of age. *Sci Transl Med* 2014;6:220ps1.
- McGlynn EA, Asch SM, Adams J, Keesey J, Hicks J, De-Cristofaro A, et al. The quality of health care delivered to adults in the United States. *N Engl J Med* 2003;348:2635–45.
- The Lewin Group. The value of diagnostics: innovation, adoption and diffusion into health care. <http://www.lewin.com/publications/Publication/237/> (Accessed September 2014).
- The Lewin Group. Laboratory medicine: a national status report. <http://www.lewin.com/publications/publication/343/> (Accessed September 2014).
- Epner PL. Achieving the promise of point-of-care testing. *Point Care* 2014;13:112–3.
- Bingisser R, Cairns C, Christ M, Hausfater P, Lindahl B, Mair J, et al. Cardiac troponin: a critical review of the case for point-of-care testing in the ED. *Am J Emerg Med* 2012;30:1639–49.
- Heneghan P, Alonso-Coello J, Garcia-Alamino JM, Perera R, Meats E, Glasziou P. Self-monitoring of oral anticoagulants: a systematic review and meta-analysis. *Lancet* 2006;367:404–11.
- Jang JY, Shin SD, Lee EJ, Park CB, Song KJ, Singer AJ. Use of a comprehensive metabolic panel point-of-care test to reduce length of stay in the emergency department: a randomised controlled trial. *Ann Emerg Med* 2013;61:145–51.
- Lee-Lewandowski E, Lewandowski K. Perspectives on cost and outcomes for POCT. *Clin Lab Med* 2009;29:479–89.
- Hernandez JS. Cost-effectiveness of laboratory testing. *Arch Pathol Lab Med* 2003;127:440–5.
- Fang CH, Otero HJ, Greenberg D, Neumann PJ. Cost-utility analyses of laboratory tests: a systematic review. *Value Health* 2011;14:1010–8.
- Levinson, DR. Comparing lab test payment rates: Medicare could achieve substantial savings. US Department of Health and Human Services, Office of the Inspector General 2013;OEI-07-11-00010.
- Moynihan R, Doust J, Henry D. Preventing overdiagnosis: how to stop harming the healthy. *BMJ* 2012;344:e3502.
- Epner PL, Gans JE, Graber ML. When diagnostic testing leads to harm: a new outcomes-based approach for laboratory medicine. *BMJ Qual Saf* 2013;22(Suppl 2):ii6–10.
- Astion M, Shojania KG, Hamill TR, Kim S, Ng VL. Classifying laboratory incident reports to identify problems that jeopardize patient safety. *Am J Clin Pathol* 2003;120:18–26.
- Plebani M. Detection and prevention of errors in laboratory medicine. *Ann Clin Biochem* 2010;47:101–10.
- Plebani M. Laboratory-associated and diagnostic errors: a neglected link. *Diagnosis* 2014;1:89–94.
- Miller WG, Myers GL, Gantzer ML, Kahn SE, Schonbrunner R, Thienpont LM, et al. Roadmap for harmonization of clinical laboratory measurement procedures. *Clin Chem* 2011;57:1108–17.
- Kost GJ, Hale KN. Global trends in critical values practices and their harmonization. *Clin Chem Lab Med* 2011;49:167–76.
- Bissell MG. Laboratory-related measures of patient outcomes: an introduction. Washington: AACC Press: 2000. p 194.
- Institute of Medicine. *Crossing the Quality Chasm: a new health system for the 21st century*. Washington (DC): National Academy Press: 2001.
- Moons KGM, de Groot JAH, Linnet K, Reitsma JBR, Bossuyt PMM. Quantifying the added value of a diagnostic test or marker. *Clin Chem* 2012;58:1408–17.
- Anonychuk A, Beasall G, Shorter F, Kloss-Wolf R, Neumann P. A framework for assessing the value of laboratory diagnostics. *Health Manage Forum* 2012;25:S4–11.
- Bossuyt PM, McCaffrey K. Additional patient outcomes and pathways in evaluations of testing. *Med Decis Making* 2009;29:E30–8.
- Lee DW, Neumann PJ, Rizzo JA. Understanding the medical and nonmedical value of diagnostic testing. *Value Health* 2010;13:310–4.
- Hickner J, Thompson PJ, Wilkinson T, Epner P, Sheehan M, Pollock AM, et al. Primary care physicians' challenges in ordering clinical laboratory tests and interpreting results. *J Am Board Fam Med* 2014;27:268–74.
- Epner P, Astion M. Focusing on test ordering practices to cut diagnostic errors. *Clin Lab News* 2012;38(7):17–18.
- Plebani M. Interpretative commenting: a tool for improving the laboratory/clinician interface. *Clin Chim Acta* 2009;404:46–51.
- Laposata M. Errors in clinical laboratory test selection and result interpretation: commonly unrecognised mistakes as a cause of poor patient outcome. *Diagnosis* 2014;1:85–7.
- Zhi M, Ding EL, Theisen-Toupal J, Whelan J, Arnaut R. The landscape of inappropriate laboratory testing: A 15-year meta-analysis. *PLoS ONE* 2013;8:e78962.
- Groopman J. *How doctors think*. New York: Houghton Mifflin; 2007.
- Gandhi TK, Kachalia A, Thomas EJ, Puopolo AL, Yoon C, Brennan TA, et al. Missed and delayed diagnoses in the ambulatory setting: a study of closed malpractice claims. *Ann Intern Med* 2006;145:488–96.
- Black WC. Advances in radiology and the real versus apparent effects of early diagnosis. *Eur J Radiol* 1998;27:116–22.
- McGregor MJ, Martin D. Testing 1,2,3: is overtesting undermining patient and system health? *Can Fam Physician* 2012;58:1191–8.
- Wu AHB, Bolger AF, Hollander JE. Growing pains with the use of high-sensitivity cardiac troponin assays. *J Am Coll Cardiol* 2013;62:1250–1.
- Hoerger TJ, Wittenborn JS, Young W. A cost-benefit analysis of lipid standardization in the United States. *Prev Chronic Dis* 2011;8:A136.
- World Alliance for Patient Safety. Summary of the evidence on patient safety: implications for research. Geneva; World Health Organization: 2008.
- Callen J, Georgiou A, Li J, Westbrook JI. The safety implications of missed test results for hospitalised patients: a systematic review. *BMJ Qual Saf* 2011;20:194–9.
- Kilpatrick ES, Holding S. Use of computer terminals on wards to access emergency test results: a retrospective audit. *BMJ* 2001;322:1101–3.
- Singh H, Spitzmueller C, Petersen N, Sawhney MK, Sittig DF. Information overload and missed test results in electronic health record-based settings. *JAMA Intern Med* 2013;173:702–4.
- Turner HE, Peake RWA, Allison JJ. The impact of centrifugation in primary care on pseudohypokalaemia: a retrospective evaluation. *Ann Clin Biochem* 2013;50:371–3.
- O'Kane MJ. The evidence base for laboratory medicine: more work needed. *Ann Clin Biochem* 2013;50:293–5.
- Bruns DE, Boyd JC. Assessing the impact of biomarkers on patient outcome: an obligatory step. *Scand J Clin Lab Invest* 2010;70(Suppl 242):85–9.
- Neumann PL, Tunis SR. Medicare and medical technology: the growing demand for relevant outcomes. *N Engl J Med* 2010;362:377–9.
- Whiting P, Rutjes AW, Reitsma JB, Glas AS, Bossuyt PM, Kleijnen J. Sources of variation and bias in studies of diagnostic accuracy: a systematic review. *Ann Intern Med* 2004;140:189–202.
- Price CP. Evidence-based laboratory medicine: is it working in practice? *Clin Biochem Rev* 2012;33:13–9.
- Price CP, Christenson RH. Ask the right question: a critical step for practicing evidence-based laboratory medicine. *Ann Clin Biochem* 2013;50:306–14.
- Horvath AR, Lord SJ, StJohn A, Sandberg S, Cobbaert C, Lorenz S, et al. From biomarkers to medical tests: the changing landscape of test evaluation. *Clin Chim Acta* 2014;427:49–57.
- Bossuyt PM, Reitsma JB, Linnet K, Moons KG. Beyond diagnostic accuracy: the clinical utility of diagnostic

- tests. *Clin Chem* 2012;58:1636–43.
59. Mallett S, Halligan S, Thompson M, Collins GS, Altman DG. Interpreting diagnostic accuracy studies for patient care. *BMJ* 2012;344:e3999.
 60. Lewin Group. Laboratory medicine and comparative effectiveness research. <http://www.lewin.com/publications/publication/402/> (Accessed September 2014).
 61. Lord SJ, Irwig I, Bossuyt PM. Using the principles of randomized controlled trial design to guide test evaluation. *Med Decis Making* 2009;29:E1–12.
 62. Trikalinos TA, Siebert U, Lau J. Decision-analytic modeling to evaluate benefits and harms of medical tests: uses and limitations. *Med Decis Making* 2009;29:E22–29.
 63. Bossuyt PM, Reitsma JB, Bruns DE, Garsonis CA, Glasziou PP, Irwig LM, et al. Towards complete and accurate reporting of studies of diagnostic accuracy: the STARD initiative. *Clin Chem* 2003;49:1–6.
 64. Schultz KF, Altman DG, Moher D; CONSORT Group. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. *BMC Medicine* 2010;8:18.
 65. Agency for Healthcare Research and Quality. Methods guide for medical test reviews. Publication No. 12-EC017. Rockville; Agency for Healthcare Research and Quality: 2012. <http://effectivehealthcare.ahrq.gov/search-for-guides-reviews-and-reports/?pageaction=displayproduct&productid=558> (Accessed September 2014).
 66. Whiting P, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med* 2011;155:529–36.
 67. Ferrante di Ruffano L, Hyde CJ, McCaffery KJ, Bossuyt PM, Deeks JJ. Assessing the value of diagnostic tests: a framework for designing and evaluating trials. *BMJ* 2012;344:e686.
 68. Institute of Medicine of the National Academies. Finding what works in health care: standards for systematic reviews. <http://iom.edu/Reports/2011/Finding-What-Works-in-Health-Care-Standards-for-Systematic-Reviews.aspx> (Accessed September 2014).
 69. Simera I, Altman DG. Writing a research article that is “fit for purpose”: EQUATOR Network and reporting guidelines. *Evid Based Med* 2009;14:132–4.
 70. Soletormos G, Duffy MJ, Hayes DF, Sturgeon CM, Barak V, Bossuyt PM, et al. Design of tumor biomarker-monitoring trials: a proposal by the European Group on Tumor Markers. *Clin Chem* 2013;59:52–9.
 71. Korevaar DA, Ochodo EA, Bossuyt PM, Hooft L. Publication and reporting of test accuracy studies registered in ClinicalTrials.gov. *Clin Chem* 2014;60:651–9.
 72. Christenson RH, Snyder SR, Shaw CS, Derzon JH, Black RS, Mass D, et al. Laboratory medicine best practices: systematic evidence review and evaluation methods for quality improvement. *Clin Chem* 2011;57:816–25.
 73. Kilpatrick ES, Freedman DB; National Clinical Biochemistry Audit Group. A national survey of interpretative reporting in the UK. *Ann Clin Biochem* 2011;48:317–20.
 74. Kilpatrick ES, Rigby AS, Atkin SL, Barth JH. Glycemic control in the 12 months following a change to SI hemoglobin A1c reporting units. *Clin Chem* 2013;59:1457–60.
 75. Misra S, Barth JH. Guidelines are written, but are they followed? *Ann Clin Biochem* 2013;50:400–2.
 76. Beall GH. Adding value to laboratory medicine: a professional responsibility. *Clin Chem Lab Med* 2013;51:221–7.