Opinion Paper

Tommaso Trenti*, Holger J. Schünemann and Mario Plebani

Developing GRADE outcome-based recommendations about diagnostic tests: a key role in laboratory medicine policies

DOI 10.1515/cclm-2015-0867
Received September 5, 2015; accepted September 11, 2015; previously published online October 10, 2015

Abstract: Harmonisation and risk management policies represent key-issues in modern laboratory medicine as they focus on a more patient-centred delivery of laboratory information based on the recognition of the importance of all steps of the total testing process (TTP) for assuring quality and patient safety. However, a further essential step in project aiming to improve the value of laboratory medicine becomes the assessment of the impact of testing on patient-important outcomes. The grading of recommendations assessment, development and evaluation (GRADE) evidence to decision (EtD) frameworks may provide a systematic and transparent approach for translating the best clinical evidence available into healthcare decisions and recommendations. GRADE is a tool appropriate not only for evaluating test accuracy but also for clinical impact, such as mortality, morbidity, symptoms, and quality of life and therefore it should be applied to the outcome research in laboratory medicine. The application of GRADE requires the recognition that a recommendation about the use of test results should result from a balance between the desirable and the undesirable consequences, including non-health related consequences such as resource utilisation, feasibility, acceptability, equity and other factors. GRADE EtDs, represents a fundamental step in projects designed to improve care quality. Patient-physician-laboratory feedback can be assured through the GRADE process, where the team developing the recommendations should include the "three-parties" representatives; clinicians, laboratorians and patient/consumers. This ensures that the laboratory-patient interaction should not be a one-way process only (information from laboratory to patient) but a two-way process, incorporating patient expectations and feedback.

Keywords: GRADE; harmonisation; risk management.

***Corresponding author: Tommaso Trenti,** Laboratory Medicine Department, Ospedale Civile Sant'Agostino Estense, Modena, Via Pietro Giardini 1355, 41126, Modena, Italy, Phone: +390593961467, E-mail: t.trenti@ausl.mo.it
Holger J. Schünemann: Department of Clinical Epidemiology and Biostatistics, McMaster University Health Sciences Centre, Main Street West, Hamilton, ON L8S 4K1, Canada
Mario Plebani: Department of Laboratory Medicine, University-Hospital of Padova, Via Giustiniani 2, Padova, Italy.
http://orcid.org/0000-0002-0270-1711

Introduction

Harmonisation

Harmonisation [1–3] and risk management policies [4, 5] need to be integrated with the right utilisation of laboratory tests that improve patient outcomes. According to the Clinical and Laboratory Standards Institute (CLSI) definition, harmonisation is "the process of recognising, understanding, and explaining differences while taking steps to achieve uniformity of results, or at minimum, a means of conversion of results such that different groups can use the data obtained from assays interchangeably" [6]. However, a broader view of harmonisation should take into consideration not only "results", but also all steps that affect laboratory information, namely the appropriateness in test request, interpretation and utilisation in the diagnostic-therapeutic process. In fact, what counts for the right patient management, and that should be, therefore, harmonised, is the laboratory information which is significantly affected by the quality of the pre- (e.g. appropriate request, quality of the biological sample, etc.) and post-analytical phase (e.g. appropriate reference ranges, interpretative comments, valuable turnaround time, etc.).

There is awareness that the lack of harmonisation in test results, names, units, reference intervals, commutability and acceptable degrees of uncertainty are not only cause for confusion, but may also be potentially

dangerous. The cases of cardiac troponin I and/or T [7] and the unit for haemoglobin expression are well-known examples of how patient safety could potentially be affected [8].

The term "harmonisation" is generally intended to guarantee that test results are equivalent, being either traceable to a reference material or based on a consensus approach in agreement with the mean values obtained with different methods [9–11]. The concepts of commutability, uncertainty and reference intervals to harmonise laboratory results are well known issues in laboratory policy to achieve harmonisation. A growing body of evidence demonstrates that clinical benefits can be achieved only by focusing on the total testing process (TTP), and, in particular, on the appropriateness of test requesting and interpretation. However, if the scope of harmonisation goes beyond method and analytical results to consider all other aspects of laboratory testing, including strategies for test demand and criteria for result interpretation [9, 10], the cooperation at the clinical-laboratory interface becomes essential to guarantee a valuable medical decision-making process and effective patient care based on outcome evaluation. However, the outcome or the improvement of the patients' health remains a "holy grail" as the assessment of the effect of the test on patient outcomes is very difficult to perform.

Risk management

The Institute of Medicine (IOM) report, "To err is human" considerably amplified the degree of concern around adverse events and patient safety in healthcare, including errors in laboratory medicine [12]. In this light, new approaches to quality and patient safety in the health care system emphasise that diagnostic process improvements should be based on assuring desired outcomes rather than focusing on the identification of unusual errors. The outcome-based approach recently proposed by Epner et al. [13] on testing-related diagnostic errors calls for a more effective selection (pre-pre) and interpretation (post-post) of clinically useful biomarkers in order to prevent adverse events, failure to diagnose and appropriately treat. Patient safety is compromised not only by inappropriately requested tests but also by misinterpretation of the results. Thus, the relevance of harmonisation and risk management policies in the laboratory increasingly recognises the need to consider patient outcomes in the assessment of tests and test strategies. The use of tests should be guided by reliable, evidence-based recommendations incorporating a multi-professional contribution and an outcome-based approach [4, 9]. A further driver to promote harmonisation and risk management is evaluation of proposed new biomarkers and in general, innovative diagnostic tests that need to improve the framework for the evaluation and recommendations models, in order to be appropriate and effective in their use.

The challenges

Laboratory tests are currently not perceived as playing a primary role in providing added value to medical treatments, as diagnostic tests are often considered a "commodity" in health resource [14, 15]. A great challenge in laboratory medicine is communicating to users about the best biomarker that contributes to the best downstream clinical pathway, including management based on test results. There is poor quality and limited evidence to prove the pivotal role of laboratory testing in value-added medicine, however some observational studies and clinical audit projects showed the essential role of testing in the global clinical, and patient management processes [16, 17]. The key issues are described in Table 1.

The challenge is the focus on patient-important outcomes and utilising the best clinical evidence related to the use of diagnostic tests in the context of harmonisation and laboratory risk management policies. The grading of recommendations assessment, development and evaluation (GRADE) approach to assess the certainty in evidence (also known as quality of evidence, strength of evidence or confidence in estimates) and develop recommendations, thus facilitating decisions in health care, is a widely, patient-centred approach [18]. This approach is currently utilised by over 90 organisations worldwide and has become the standard for providing health care recommendations [19].

Appropriateness of test requests

Inappropriate test demand is defined as "a test request that is made outside some form of agreed guidance" [20]. The focus on "guidance" reflects the scientific and cultural concept of "evidence-based medicine" which emphasises that the best available evidence should "guide" health care practice. However, to date there is scarce evidence regarding the impact of unnecessary test requests, because most dedicated studies do not meet methodological acceptable standards due to bias in study design, including the lack of consensus for the definition of an "inappropriate test request" [21, 22].

Table 1: Key issues in an outcome-oriented view of harmonisation and risk management projects in laboratory medicine.

- In the context of harmonisation and laboratory risk management policies there is the need for evaluating the effects of diagnostic tests as patient-important outcomes
- Tests should be considered in terms of their specific use and the resulting downstream clinical actions
- The purpose and key outcomes of a test should be clearly defined in order to develop a completely effective decision making model for diagnostic intervention
- The diagnostic test process requires validation and interaction of all the steps in the total testing process to achieve harmonisation
- To provide the best information for decision-making for a clinical question, traditional analytical specifications should guarantee clinical diagnostic performance while acknowledging that they are only part in the test value evaluation
- The grading of recommendations assessment, development and evaluation (GRADE) evidence to decision (EtD) frameworks may provide a systematic and transparent approach for translating the best clinical evidence available into healthcare decisions and recommendations
- Laboratory-patient/consumer interaction should constitute a two-way process, based on a policy that is actively promoted by laboratory professionals and aimed at informing patient/consumer about the required diagnostic test
- Patient/consumer information should be based on outcome evaluations as the core value, and GRADE seems a valid working strategy due to its robust scientific and transparent methodological approach

The role of quality improvement projects

To achieve harmonisation in test request and results interpretation and utilisation, laboratory medicine should provide advice to physicians in their selection and interpretation. The decisions about test request and interpretation are however complex and Evidence Based Medicine (EBM) may be an effective tool to allow this goal. Table 2 visually presents the similarities between the outcome-based approach to testing-related diagnostic errors and harmonisation in laboratory medicine: both ideally incorporating an outcome evaluation to guide the decision-making process.

The GRADE approach should increase awareness of laboratory professionals in harmonisation projects and risk management policies with its focus on patient-important outcomes. Appropriate guidelines in influencing clinical practice depends on the quality, acceptance and implementation and effective communication strategies should help to bridge the gap between clinical research and currently healthcare practice [23].

Assessing outcomes as a result of laboratory testing

As recently proposed by Barth, "a quality clinical service laboratory might be simply described as performing the right test on the right person at the right time and interpreting that test correctly" [24]. If harmonisation goes beyond analytical methods and strategies, and risk management extends beyond the realms of the laboratory, the ultimate goal becomes the assessment of the impact of testing on patient-important outcomes. The key question therefore is how to evaluate test use and application in terms of these outcomes, and how to assess utility and effectiveness, with an appropriate methodology, strong

Table 2: Relationships between laboratory harmonisation in the total testing process as proposed by Plebani and Panteghini [9], and the outcome-based approach to testing-related diagnostic errors as proposed by Epner et al. [13].

Outcome-based approach to testing-related diagnostic errors	Harmonisation in Laboratory Medicine (in the total testing process)
Source: Processes external to the laboratory – An inappropriate test is ordered – An appropriate test is not ordered – An appropriate test result is misapplied	Initial and/or final steps of the total diagnostic process (outside the laboratory) – Selection of reference biomarkers – Appropriateness in test request – Appropriate test interpretation and decision to be acted – References population data base
Source: Internal processes (within the laboratory) – Accuracy and reproducibility of analytical results (Internal quality control and external quality assessment schemes) – An appropriate test is ordered, but a delay occurs somewhere in the total testing process – The result of an appropriately ordered test is inaccurate	Internal processes (within the laboratory) – Accuracy and reproducibility of analytical results (Internal quality control and external quality assessment schemes) – Evaluation of pre-analytical sources and pre- analytical quality – Harmonisation of currently available assays and analytical control practice

enough to develop high quality guidance. In the context of harmonisation this requires an agreement on the strategies to improve test requests, related analytical performances and the assurance that the interpretation of laboratory results is correct and utilised appropriately in patient care [25].

The GRADE (the grading of recommendations assessment, development and evaluation) approach

The GRADE working group developed a transparent method for grading the quality of research evidence and strength of recommendations to guide health care practice. GRADE is a tool appropriate not only for evaluating test accuracy but also on clinical impact, such as mortality, morbidity, symptoms, and quality of life [26, 27]. Analytical and diagnostic performances or accuracy such as sensitivity, specificity, imprecision, positive predictive value (PPV) and negative predictive value (NPV) are traditionally established measures of test accuracy. In addition, although diagnostic testing recommendations share the fundamental logic of treatment recommendations, they present unique challenges. Sensitivity, specificity, PPV, NPV, likelihood ratios, and diagnostic odds ratios, all measures of test accuracy, are among the challenging terms that diagnostic studies typically deliver to clinicians. Not only do clinicians have difficulties remembering the definitions and calculations for these terms, these concepts are often complex to apply to individual patients. The clinical impact and health care outcomes to which these accuracy measures relate and which should be the final goal of the process, are more complex to measure and evaluate, and are therefore often not considered. GRADE places emphasis on relating these accuracy measures, as surrogates, to patient-important outcomes [28]. The GRADE methodology has been extensively used for grading the quality of evidence and strength of recommendations for therapeutic questions [29], prognosis and is increasingly used in the area of medical testing [30–34]. The GRADE framework is turning away from simple test accuracy to incorporate main health outcomes in the thinking process about best use of tests. Direct studies assessing the impact of diagnostic tests or strategies on patient important outcomes are rarely available. In the absences of such evidence, the GRADE process requires, following the development of health care questions that address patient important outcomes and assessing the confidence in test accuracy data, two main steps. The former is the judgments about directness concerned in assessing the link between test accuracy and important health outcomes, and the latter aims on the criteria used in moving from evidence to a recommendation or decision for use of diagnostic tests in suggested strategies [35].

The application of GRADE requires the recognition that a recommendation about the use of test results from a balance between the desirable and the undesirable consequences, including non-health related consequences such as resource utilisation, feasibility, acceptability, equity and other factors [36, 37]. GRADE, in the context of making recommendations or decisions about tests, disintegrates the steps that are required for a proper evaluation of the linked pieces of evidence going from testing to other management. This evaluation is best achieved in the context of multi-disciplinary panels. However, the understanding about diagnostic tests by members of the guideline panels and the methodology for developing recommendations is far from being completely explored [38, 39]. The GRADE approach offers structure and transparency in the complex process of making evidence-based recommendations and has developed evidence to decision (EtD) frameworks to achieve this.

The GRADE evidence to decision (EtD) frameworks

In order to assess or model the consequences of a decision about a test and assess the certainty of the evidence, EtD frameworks for tests can be used. The frameworks include traditional criteria to assess test accuracy but also include assessment of the certainty of the additional different types of evidence used to estimate the effects of tests on final patient outcomes. A clear clinical question and outcomes (crucial to the patient) to be defined from the outset, and then a structured systematic review of the available evidence is performed. Evidence quality about diagnostic test accuracy is then assessed by considering eight criteria, five of which criteria are used to downgrade the quality of evidence, such as risk of bias, indirectness, inconsistency, imprecision, and publication bias. The remaining three criteria are used to upgrade evidence quality, such as the magnitude of the effect, dose response in relation to the effect, and opposing plausible residual bias or confounders, the GRADE evidence to decision (EtD) frameworks for tests offer a structured approach, as shown in Table 3.

Table 3: The GRADE evidence to decision (EtD) frameworks for a test structured approach.

1. Formulating the question
2. Making an assessment
– The Problem
– Test accuracy
– Benefits and harms
3. Certainty of the evidence
– What is the certainty of the evidence of test accuracy?
– What is the certainty of the evidence for any critical or important direct benefits, adverse
– What is the certainty of the evidence of effects of natural history or the management that is guided by the test results?
– How certain is the link between test results and management decisions?
– What is the overall certainty of the evidence about the effects of the test?
4. Values
5. Balance between the desirable and undesirable effects
6. Resource use
7. Equity, acceptability and feasibility

Formulating the question

Formulating a question requires a clear outline of the problem, purpose, type and role of a test, alternative intervention(s), the main outcomes and the setting. The population intervention comparison outcome (PICO) format offers a method for the formulation of the question about a diagnostic test which specifies patient important outcomes, rather than relying on test accuracy only [40].

Making an assessment

The problem

At the basis of the assessment a definition of the magnitude and priority of the problem should be established, and this depends on perspective, the setting in which the test is/will be used and the influence on current/future practices. In the case of suggested new test introduction, accuracy, adverse effects, availability, costs and limitations or substitution of currently used tests should also be considered.

Test accuracy

Interpretation of test accuracy should be based on a summary of findings from systematic reviews. A test to enter into an EtD framework evaluation, requires an acceptable overall accuracy as a starting point, otherwise the assessment should not proceed.

Benefits and harms

Findings, according to desirable and undesirable effects, form the basis of the judgments about the benefits and harms of using a test. Evidence should stem from up-to-date systematic reviews and summarised in a table of findings [41]. Detailed assumptions and calculations for transparency are at the basis of an ideal approach.

Certainty of the evidence

The overall rating of the certainty of the evidence about the effects of using a test (and subsequent management decisions) on patient-important outcomes should be based on the certainty of the evidence, considering the weakest link in the chain of evidence used to estimate those effects [42]. The GRADE approach for evaluating the certainty of the evidence of effects for clinical interventions is now widely used by guideline developers [43] and a detailed description of this approach can be found elsewhere [30]. EtD frameworks for tests include five criteria for making judgments and rating the certainty of the evidence. As reported in Table 3, the framework assists in quantifying the certainty of 1) test accuracy, (2) any critical or important direct benefit, adverse effects or burden of the test, (3) effects of natural history or the management that is guided by the test results, (4) the link between test results and management decisions and (5) the evidence about the effects of the test.

Values

The perceived value of the main outcomes and the affect of the decision, and in the case of tests, this includes test burden and the downstream outcomes, is addressed. For example, a blood test may substitute more invasive diagnostic interventions, such as bowel biopsy for coeliac disease diagnosis, prostate biopsy for tumor, or fetal cell genotyping through maternal blood sampling instead of

amniocentesis. Declarations of value uncertainty should also be included in the assessment framework.

Balance between the desirable and undesirable effects

Desirable and undesirable effects of a test need to be judged through either formal or informal modeling, which in turn effect the downstream actions from a test result interpretation. For instance, in the case of cancer biomarkers, further intervention may be evaluated in the likelihood of true or false positives, and balanced in terms of presumptive outcome.

Resource use

In the case of the selection of the proposed diagnostic test, judgments about the magnitude of costs, certainty of evidence of resource requirements and the cost-effectiveness of interventions should include the evaluation of the impact both within the laboratory and the downstream consequence. Usually the cost of performing a test is very low, but the impact of downstream action can be great, such as when a test correctly identifies the state of a patient (true positive or negative), reducing unnecessary hospitalisation, hospital stay, further diagnostic procedures and assisting in a timely and correct therapeutic response. Whereas, the cost of an ineffective test (false positive or negative) has a large impact on the downstream resource use, the great challenge is to identify the total health care cost and not only the direct cost of the test itself.

Equity, acceptability and feasibility

For tests, assessments of equity, acceptability and feasibility include consideration of both the test and linked interventions in the context of the health system and stakeholders. The use of specific tests in different professional settings for the same clinical presentation, concerns equity of access to clinical care. For example in the same regional health care setting, the utilisation of a specific test may vary from one hospital to another.

A potential direct patient and laboratorian's brain loop link

The harmonisation of policies for test request, interpretation an appropriate utilisation, derived from data based on clinical outcomes and, where possible, supported by EBM, through a process such as GRADE, in particular through GRADE EtDs, represents a fundamental step in projects designed to improve care quality. Patient-physician-laboratory feedback can be assured through the GRADE process, where the team developing the recommendations should include the "three-parties" representatives; clinicians, laboratorians and patient/consumers. This ensures that the laboratory-patient interaction should not be a one-way process only (information from laboratory to patient) but a two-way process, incorporating patient expectations and feedback. Assessing the most effective methods consumer involvement in future recommendations for promoting a clinical diagnostic process oriented to the best and most effective test request therefore becomes essential.

Economic drivers of direct laboratory access may overlook important issues concerning appropriateness, analytical reliability, and effectiveness on improving clinical outcomes. This issue is particularly important in genomics and pharmacogenomics, where misleading and/or counterproductive information, could influence consumers in their complex medical decision making process without appropriate clinical information, raising concerns about the risks of inappropriate actions following the direct diagnostic test results communication to patients. The Royal College of Pathologists released an explicit document stating that several laboratory test results require "professional interpretation rather than measurement, and the laboratory interpretation may need to be modified by the clinician who knows the patient's specific situation, or even at a multidisciplinary team meeting" [44]. As the GRADE approach is now recognised and internationally used [28, 45, 46] to develop ways to present concise summaries of the findings of systematic reviews at the basis for recommendations or decisions about diagnostic tests and health outcomes [19, 47], GRADE has also addressed the presentation of recommendations to health professionals, policymakers, and patients/consumers. The Developing and Evaluating Communication Strategies to Support Informed Decisions and Practice Based on Evidence (DECIDE) project by the GRADE working group and partners [48, 49] has explored methods of effectively communicating evidence-based diagnostic test recommendations targeted individually to the key stakeholders and patients, ensuring access to concise, accurate evidence summaries to inform and assist in decision making.

In this light, laboratory societies in association with governmental institutions, through an approach such as GRADE, should be able to promote active policies of directly targeting patients with harmonised information

to optimise health prevention strategies and the utility of appropriate tests for target populations. The active contact should promote the concept that "patient empowerment" can be achieved through the correct test request with the assistance of a medical practitioner, and discouraging direct and inappropriate use of diagnostic test requests. Lab Test Online [50] should be a good example of a direct information policy designed for patients, where the content is guaranteed and harmonised by the leading national/international laboratory societies, worldwide.

Further crucial issues involve the quality of test results from non-accredited laboratories and inappropriate testing in an inappropriate population. The Prostate-specific antigen (PSA) test is a paradigmatic example of the concerns regarding inappropriate and possibly low quality testing, currently available in local pharmacies, and the increasing risk of false positives. A global picture of the testing process should be consider correct interpretation and clinical utilisation of laboratory data, including analytical standardisation with appropriate revision of reference values and decision limits. The PSA is a representative example of a diagnostic test that requires high quality in terms of both analytical performances, and an informed patient regarding the right request, interpretation and further correct actions to avoid inappropriate decisions and treatments. In this setting the theory of analytical standardisation introducing the term "reference" to replace the ambiguous, arbitrary and even misleading definition of "normal", is hard for the patients to interpret. However if analytical standardisation is not followed by the revision of reference values and decision limits, for a correct interpretation and clinical utilisation of laboratory data, patients' outcomes could be worsened, as demonstrated in three effective examples: serum creatinine and glomerular filtration rate equations, the prostate-specific antigen recalibration and the glycated haemoglobin standardisation [10, 11].

Analytical specification should guarantee the right clinical diagnostic performances in response to the clinical question to give the right information. Health outcome should be the ultimate goal and when possible, all the diagnostic steps in the harmonisation and risk management process evaluated in terms of effectiveness.

Conclusions

The need for patient outcome evaluation is a key issue in the development of harmonisation and risk management policies in Laboratory Medicine to achieve an effective Clinical Governance [51]. Behind the traditional analytical specifications able to guarantee the required diagnostic performance in response to the clinical question, the outcomes of a test should be clearly defined and evaluated. GRADE is a methodology suitable to develop an effective decision-making model, the laboratory test should be fully considered in light of the resulting downstream clinical actions [4, 9, 52]. However, this is a difficult process and health outcome should be the ultimate goal. GRADE Evidence to Decision (EtD) frameworks provide a systematic and transparent approach for translating the best clinical evidence available into healthcare decisions and recommendations, as the frameworks are providing an approach to take into account all test effects and consequences on what matters to those mostly affected, the patients.

Author contributions: All the authors have accepted responsibility for the entire content of this submitted manuscript and approved submission.
Research funding: None declared.
Employment or leadership: None declared.
Honorarium: None declared.
Competing interests: The funding organisation(s) played no role in the study design; in the collection, analysis, and interpretation of data; in the writing of the report; or in the decision to submit the report for publication.

References

1. Ambruster D, Miller RR. The Joint Committee for Traceability in Laboratory Medicine (JCTLM): a global approach to promote standardization of clinical laboratory test results. Clin Biochem Rev 2007;28:105–13.
2. Plebani M. Harmonization in laboratory medicine: the complete picture. Clin Chem Lab Med 2013;51:741–51.
3. Miller WG, Myers GL, Gantzer ML, Khan SE, Sconbrunner ER, Thienpont LM, et al. Roadmap for harmonization of clinical laboratory measurement procedures. Clin Chem 2011;57:1108–17.
4. McLawhon RW. Patient safety and clinical effectiveness as imperatives for achieving harmonization inside and outside the clinical laboratory. Clin Chem 2011;57:936–83.
5. Plebani M. The CCLM contribution to improvements in quality and patient safety. Clin Chem Lab Med 2013;51:39–46.
6. Clinical and Laboratory Standards Institute (CLSI) Harmonized Terminology Database. Available from: http://www.clsi.org/source/custom/termsall.cfm?Section=Harmonized_Terminology_Database. Accessed 2 Mar 2015.
7. Ungerer JP, Marquart L, O'Rourke PK, Wilgen U, Pretorious CJ. Concordance, variance, and outliers in 4 contemporary cardiac troponin assays: implications for harmonization. Clin Chem 2012;58:274–83.
8. De la Salle B. Pathology harmony moves on: progress on implementation in haematology. Br J Haematol 2012;366:780–1.

9. Plebani M, Panteghini M. Promoting clinical and laboratory interaction by harmonization. Clin Chim Acta 2014;432:15–21.
10. Tate JR, Johnson R, Barth J, Panteghini M. Harmonization of laboratory testing – current achievements and future strategies. Clin Chim Acta 2014;432:4–7.
11. Panteghini M. Implementation of standardization in clinical practice: not always an easy task. Clin Chem Lab Med 2012;50:1237–41.
12. Kohn LT, Corrigan JM, Donaldson MS. To err is human: building a safer health system. Washington, DC: National Academy Press, 1999.
13. Epner PL, Janet EG, Graber ML. When diagnostic testing leads to harm: a new outcomes-based approach for laboratory medicine. BMJ Qual Saf 2013;0:1–5.
14. Plebani M, Lippi G. Is laboratory a dying profession? Blessed are those who have not seen and yet have believed. Clin Biochem 2010;43:939–41.
15. Plebani M. Clinical laboratories: production industry or medical services? Clin Chem Lab Med 2015;53:995–1004.
16. Lundberg GD. The need for an outcome research agenda for clinical laboratory testing. J Am Med Assoc 1998;280:565–6.
17. Wahner-Roedeler DL, Chaliki SS, Bauer BA, Budrick JB, Bergstrom LR, Lee Mc, et al. Who makes the diagnosis? The role of clinical skills and diagnostic test results. J Eval Clin Pract 2007;13:321–5.
18. Schünemann HJ, Oxman AD, Brozek J, Glasziou P, Jaeschke R, Vist GE, et al. Grading the quality of evidence and strength of recommendations for diagnostic tests and strategies. Br Med J 2008;336:106–10.
19. GRADE Working Group: Organizations that have endorsed or that are using GRADE. Available from: http://www.gradeworkinggroup.org/society/. Accessed 29 May 2015.
20. Fryer AA, Smellie WS. Managing demand for laboratory tests: a laboratory toolkit. J Clin Pathol 2013;66:62–72.
21. Bossuyt PM. Room for improvement in national academy of clinical biochemistry laboratory medicine practice guidelines. Clin Chem 2012;58:1392–4.
22. Aakre KM, Langlois MR, Watine J, Barth JH, Baum H, Collison P, et al. Critical review of laboratory investigations in clinical practice guidelines: proposals for the description of investigation. Clin Chem Lab Med 2013;51:1217–26.
23. GRADE Working Group: Systems for grading the quality of evidence and strength of recommendations I: critical apparaisal of existing approaches. BMC Health Serv Res 2004;4:38.
24. Barth JH. Clinical quality indicators in laboratory medicine: a survey of current practice in the UK. Ann Clin Biochem 2011;48:238–40.
25. Schünemann HJ, Guyatt GH. Goodbye M(C)ID! Hello MID, where do you come from? Health Serv Res 2005;40:593–7.
26. Schünemann HJ, Oxman AD, Jan Brozek J, Glasziou P, Bossuyt P, Chang S, et al. EBM notebookGRADE: assessing the quality of evidence for diagnostic recommendations. Evid Based Med 2008;13:162–3.
27. Guyatt GH, Oxman AD, Vist GE, Kunz R, Falck-Ytter Y, Alonso-Coello P, et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. Br Med J 2008;336:924–6.
28. Schünemann HJ, Mustafa R, Brozek J. Diagnostic accuracy and linked evidence – testing the chain. Z Evid Fortbild Qual Gesundhwes 2012;106:153–60.
29. Balshem H, Helfand M, Schünemann HJ, Oxman AD, Kunz R, Brozek J, et al. GRADE guidelines: 3. Rating the quality of evidence. J Clin Epidemiol 2011;64:401–6.
30. World Health Organisation. Policy statement: automated real-time nucleic acid amplification technology for rapid and simultaneous detection of tuberculosis and rifampicin resistance: Xpert MTB/RIF system. Geneva, Switzerland: World Health Organization, 2011.
31. Hsu L, Brozek JL, Terracciano L, Kreis J, Compalati E, Stein AT, et al. Application of GRADE: making evidence-based recommendations about diagnostic tests in clinical practice guidelines. Implement Sci 2011;6:62.
32. Guyatt GH, Akl EA, Crowther M, Schünemann HJ, Gutterman DD, Zelman LS. Introduction to the ninth edition: antithrombotic therapy and prevention of thrombosis, 9th ed.: American College of Chest Physicians Evidence-Based Clinical Practice Guidelines. Chest 2012;141:48S–52S.
33. Fiocchi A, Schünemann HJ, Brozek J, Restani P, Beyer K, Troncone R, et al. Diagnosis and rationale for action against cow's milk allergy (DRACMA): a summary report. J Allergy Clin Immunol 2010;126:1119–28.
34. Bates SM, Jaeschke R, Stevens SM, Goodacre S, Wells PS, Stevenson MD, et al. Diagnosis of DVT: antithrombotic therapy and prevention of thrombosis, 9th ed.: American College of Chest Physicians Evidence-Based Clinical Practice Guidelines. Chest 2012;141:e351S–418S.
35. Gopalakrishna G, Mustafa RA, Davenport C, Scholten RJ, Hyde C, Brozek J, et al. Applying grading of recommendations assessment, development and evaluation (GRADE) to diagnostic tests was challenging but doable. J Clin Epidemiol 2014;67:760–8.
36. Brunetti M, Pregno S, Schünemann HJ, Plebani M, Trenti T. Economic evidence in decision-making process in laboratory medicine. Clin Chem Lab Med 2011;49:617–21.
37. Brunetti M, Shemilt I, Pregno S, Vale L, Oxman A, Lord J, et al. GRADE guidelines: 10. Considering resource use and rating the quality of economic evidence. J Clin Epidemiol 2013;66:140–50.
38. Leeflang MM, JDeeks JJ, Gatsonis C, Bossuyt PM. Systematic reviews of diagnostic test accuracy. Ann Intern Med 2008;148:889–97.
39. Whiting PF, 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.
40. Harris RP, Helfand M, Woolf SH, Lohr KN, Mulrow CD, Teutsch SM, et al. Current methods of the US Preventive Services Task Force: a review of the process. Am J Prev Med 2001;20:21–35.
41. Guyatt G, Oxman AD, Akl EA, Kunz R, Vist G, Brozek J, et al. GRADE guidelines: 1. Introduction-GRADE evidence profiles and summary of findings tables. J Clin Epidemiol 2011;64:383–94.
42. Spencer FA, Iorio A, You J, Murad MH, Schünemann HJ, Vandvik PO, et al. Uncertainties in baseline risk estimates and confidence in treatment effects. Br Med J 2012;345:e7401.
43. Guyatt GH, Oxman AD, Kunz R, Vist GE, Falck-Yetter Y, Schünemann HJ, et al. What is "quality of evidence" and why is it important to clinicians? Br Med J 2008;336:995–8.
44. The Royal College of Pathologists. Statement of the Royal College of Pathologists on the delivery of medical laboratory test results direct to patients. July 30, 2010. Available from: http://

www.rcpath.org/resources/pdf/rcpath_results_direct_statement_v12.pdf. Accessed 5 Aug 2015.
45. Calonge N. New promise for uniform evidence-based guideline development: the GRADE approach. National Guideline Clearing House Expert Commentaries. 2009 Available from: http://www.guideline.gov/expert/expert-commentary.aspx?id=16440&search=grade+and+promise. Accessed on 8 august 2015.
46. Owens DK, Lohr KN, Atkins D, Treadwell JR, Reston JT, Bass EB, et al. Grading the strength of a body of evidence when comparing medical interventions. Agency for Healthcare Research and Quality and the Effective Health Care Program. J Clin Epidemiol 2010;63:513–23.
47. SUPPORT Collaboration: SUPPORT structured summaries of systematic reviews. Available from: http://www.support-collaboration.org/summaries.htm. Accessed 20 Dec 2014.
48. Treweek S, Oxman AD, Alderson P, Bossuyt PM, Brandt L, Brozek J, et al. Developing and evaluating communication strategies to support informed decisions and practice based on evidence (DECIDE): protocol and preliminary results. Implement Sci 2013;8:6.
49. Schünemann HJ, Best D, Vist G, Oxman AD. Letters, numbers, symbols and words: how to communicate grades of evidence and recommendations. CMAJ 2003;169:677–80.
50. Lab Test OnLine Available from: https://labtestsonline.org/. Accessed 1 Aug 2015.
51. Trenti T, Canali C, Scognamiglio AM. Clinical Governance and evidence-based laboratory medicine. Clin Chem Lab Med 2006;44:724–32.
52. Lippi G, Mattiuzzi C. The biomarker paradigm: between diagnostic efficiency and clinical efficacy. Pol Arch med Wewn 2015;125:208–8.