

Are guidelines guiding us on how to utilize laboratory tests?

Guest Editor: Andrea R. Horvath

SEALS Department of Clinical Chemistry and Endocrinology, Prince of Wales Hospital and School of Medical Sciences, University of New South Wales, Australia

Screening and Test Evaluation Program, School of Public Health, University of Sydney, Australia

ARTICLE INFO

Corresponding author:

Andrea Rita Horvath
SEALS North, Department of Clinical
Chemistry and Endocrinology
Level 4, Campus Centre
Prince of Wales Hospital
Barker Street, Randwick
NSW 2031, Sydney, Australia
Phone: (+61) 404 027 843
E-mail: rita.horvath@sesiahs.health.nsw.gov.au

Key words:

evidence-based laboratory medicine,
guideline, outcome, test utilisation

Disclaimer:

This article is based on a lecture given
at IFCC Worldlab 2014 in Istanbul.

ABSTRACT

Increasing patient risks and costs associated with the delivery of health care services have been related to inappropriate and uncontrolled use of biomarkers which make evidence-based guideline recommendations for best practice increasingly important. The translation of basic scientific discoveries into clinically meaningful studies and then to evidence-based clinical practice guidelines (CPGs) or health policy is, however, not straightforward. CPGs are potentially the most influential publications as they aim to guide clinical decisions and impact patient outcomes; hence, current approaches to their development often fail scientific publication standards. Critical appraisal of CPGs has revealed that many do not involve laboratory professionals in formulating recommendations on the use of tests; the composition of the panel could influence the scope of guidelines and over-represent certain stakeholders' views; numerous CPGs do not have rigorous evidence-based methodology and miss essential information important for the correct interpretation and application of laboratory results.

Numerous CPGs are released on similar topics worldwide, but their quality and content validity are highly variable and their recommendations may differ even when using the same sources of evidence. This can be due to the limitations of the evidence base, or to the lack of agreed test evaluation methods and easy-to-use evidence rating schemes that could be universally adapted to diagnostic recommendations. Furthermore, value-based judgments on the balance between benefits, harms, risks, patients' preferences and the organizational and financial aspects of care may differ among countries and regions. Addressing these issues requires careful discussions and consensus between relevant multidisciplinary stakeholders involved in the diagnosis and management of health conditions.

INTRODUCTION

Increasing patient risks and costs associated with the delivery of health care services have been related to inappropriate and uncontrolled use of both diagnostic and therapeutic interventions which make evidence-based guideline recommendations for best clinical practice increasingly important. According to the BEACH study in Australia, general practitioners have the greatest difficulty with test ordering and test interpretations for conditions/symptoms that are vague, and/or where there are no guidelines or decision support systems to guide their practice. The least difficulty was reported for conditions such as diabetes, lipids, urinary tract infections where clear management recommendations have been available (1). In response to these needs numerous clinical practice guidelines (CPGs) are released on similar topics worldwide. However, the translation of basic scientific discoveries into clinically meaningful studies and then distillation of study findings into evidence-based

practice recommendations or health policy are not straightforward and pose many methodological and implementation challenges.

WHY DO WE NEED GUIDELINES?

Guidelines are systematically developed statements that assist health care professionals and patients in making decisions about *appropriate* health care in specific clinical circumstances (2). Guidelines aim to:

- disseminate best practice based on scientific evidence;
- decrease practice variation and the potential or frequency of professional misconduct;
- improve patient safety;
- improve the quality and effectiveness of care;
- improve cost-effectiveness of care;
- facilitate training, education and continuous professional development;
- increase explicitness, transparency, patient information and autonomy of choice.

In the context of laboratory medicine, guidelines aim to improve the *appropriateness* of test utilization (i.e. test requesting and interpretation) by (3):

- promoting the use of new tests if evidence proves their **efficacy and effectiveness** – *start starting or stop stopping*
- eliminating poor or useless tests before they become widely available – *stop starting*
- removing old tests with no proven benefit from practice – *start stopping (adapted from 4)*.

Appropriateness in this context refers to care that results in more benefits than harms at

reasonable costs. For example, we have strong evidence from randomised controlled trials that screening with either faecal occult blood testing (FOBT) or sigmoidoscopy decreases the mortality of colorectal cancer (CRC) by 14%-16% and if the cancer is detected at an early localized stage, the 5-year survival rate is 90% (5). As a result, evidence-based recommendations have been issued by a number of guideline organisations and national screening programs have been initiated in many developed countries. For example the United States Preventive Services Task Force (USPSTF) recommends 1) high-sensitivity faecal occult blood testing annually, 2) colonoscopy every 10 years, or 3) sigmoidoscopy every 5 years with FOBT every 3 years for the prevention or early detection of CRC among adults aged 50–75 years (6). Due to these recommendations, the percentage of the U.S. population compliant with recommended CRC screening increased from 54% in 2002 to 65% in 2010 and stayed at the same rate by 2012, primarily through increased use of colonoscopy. To further improve clinical outcomes through the uptake of CRC screening, the CDC introduced more aggressive population-based strategies and set the target for 2014 at 80% (6). The European guidelines issued in 2013 still consider sigmoidoscopy and colonoscopy as a supplement or alternative for CRC screening (5). In the European Union (EU) in 2007 the Council Recommendation for CRC screening targeted approximately 136 million women and men in the age group of 50-74 years primarily by FOBT testing. In 2007 less than 10% of the targeted EU population (approximately 12 million) has taken part in CRC screening and 94% of those were tested by FOBT and the rest by flexible sigmoidoscopy or endoscopy (7).

McDowell et al. published a systematic review of 19 hypertension guidelines, issued between

2001 and 2011 with recommendations for monitoring for adverse drug reactions using biochemical tests in patients taking antihypertensive treatment. They found that guidelines were lacking any evidence behind advice on frequency of biochemical monitoring and both the instructions for monitoring and the extent of advice for subsequent action differed greatly and that such poorly specified recommendations were challenging for clinicians to apply in clinical practice (8).

Clinicians face even more challenges when guideline recommendations are not just vague or diverse but even conflicting. Examples of such confusion are conflicting recommendations for PSA screening from different professional organisations; e.g. in USA the USPSTF recommends against PSA screening to detect prostate cancer, whilst the American Cancer Society and the American Urological Association and many other European cancer societies recommend that patients willing to be screened discuss their options with their physician (9). Another recent area of controversy is related to the screening, diagnosis and management of gestational diabetes including debates about the merits of screening *versus* no screening, universal *versus* selective screening of high risk cases, timing and methods and cut-off glucose values used for defining the condition, and long-term management options for those who have the diagnosis (10,11). These examples illustrate the diversity and complexity of guideline development and implementation even when the same evidence base is available to guide best clinical practice and national policy.

VARIATIONS AND DIVERSITY IN GUIDELINE DEVELOPMENT AND IMPLEMENTATION PRACTICES

Numerous studies have demonstrated that the quality and content validity of guidelines are

highly variable (12-15). This is particularly true for diagnostic recommendations where the evidence base is more limited than in the field of therapeutics. These shortcomings are due to large variations in the analytical and clinical performance of laboratory methods for the same analytes, the lack of agreed test evaluation methods and easy-to-use evidence rating schemes that could be universally adapted to diagnostic recommendations. A recent review identified 12 evidence grading systems that addressed diagnostic testing. Out of these, 5 systems provided varying degree of coverage of the essential items for evidence gathering, review, assessment and linkage to recommendations. However, no single system covers all aspects and supports guideline developers in rating the strength of evidence behind recommendations for the use of laboratory tests (16). To add to the complexity, value-based judgments on the balance between benefits, harms, risks, patients' preferences and the organizational and financial or resource aspects of care may differ among countries and regions and therefore could influence the final recommendation and its grading. Addressing these issues requires a transparent, well-structured and documented process including careful discussions and consensus between relevant multidisciplinary stakeholders involved in the diagnosis and management of health conditions (for more details on grading, see the paper by Don-Wauchope et al., in this issue).

Clinical practice guidelines are potentially the most influential publications as they aim to guide clinical decisions and impact patient outcomes; hence, current approaches to CPG development are often non-systematic, lack clear organisational structure or legislative background and fail the methodological rigour of scientific publication standards. The most widely used critical appraisal tool for assessing the

methodological quality of CPGs is the Appraisal of Guidelines for Research and Evaluation (AGREE) Instrument (17). Numerous studies using the AGREE tool on various guideline topics issued by various organisations pointed to significant inconsistencies in terms of best practice recommendations provided to clinicians across the globe which may have an impact on the quality of care provided to patients. Moreover, the findings consistently showed that the least well-addressed domains within the AGREE tool were stakeholder involvement, rigour of development, applicability and editorial independence of the guideline development process (13-15,18). Critical appraisal by the AGREE tool of CPGs primarily addressing laboratory testing in various conditions has similarly revealed that many do not involve laboratory professionals; the composition of the panel could influence the scope of guidelines and over-represent certain stakeholders' views; and miss essential information important for the correct interpretation and application of test results (13,15).

The above factors easily explain why the European Observatory on Health Systems and Policies in its 2013 report also found divergent national guideline development and implementation programs in the EU (18). In addition to the quality of guideline methods, this report investigated the organisational and regulatory framework, the implementation and impact of guidelines developed for chronic non-communicable diseases such as coronary heart disease, chronic obstructive pulmonary disease, asthma, type 2 diabetes mellitus, osteoarthritis, breast cancer, cervical cancer, colorectal cancer and depressive disorders that are responsible for 70-80% of health care costs in the EU. Key findings of this report are listed below (18):

- Regulatory frameworks exist in most EU states for clinical guideline use but relevant laws are not always implemented.

- There is no obvious link between the availability of legislative frameworks and the quality and impact of guideline programs.
- Guidelines are usually developed by government and professional organizations or adopted/adapted from external sources.
- The engagement of multidisciplinary stakeholders in guideline development varies but patients and users of health services are rarely involved in the development of CPGs.
- Few organizations have quality control processes for their guidelines but if they do, they often use the AGREE instrument.

DO WE NEED GUIDELINES FOR MAKING GUIDELINES?

The mentioned shortcomings of guideline development programs are not unique to Europe and call for guidelines for developing guidelines and an assessment of the internal and external validity of recommendations before their implementation is attempted. Numerous government organisations issuing CPGs have guideline development manuals. The so-called GRADE and DECIDE project group systematically reviewed the available guideline development resources and assembled a checklist with 18 topics and 146 items in order to facilitate the standardisation of all stages of the guideline development process. The group provides an interactive webpage (<http://cebgrade.mcmaster.ca/guidecheck.html>) with links to training materials and resources for applying the checklist items (19).

The Institute of Medicine (IOM) has also issued a report entitled “Clinical practice guidelines we can trust”, in order to provide a set of standards for ensuring that guidelines present trustworthy and implementable recommendations (20). Table 1 summarises the IOM standards and their relevance to guideline development on laboratory testing.

In the field of laboratory medicine, the National Academy of Clinical Biochemistry (NACB) of the American Association of Clinical Chemistry (AACC) is a well-recognised source of guidelines. The NACB has recently updated its standard operating procedure for developing laboratory medicine practice guidelines that are based on more systematically gathered evidence (for more details see the paper by Kahn et al., in this issue). Table 2 summarises the main sources of guideline development tools that are relevant to laboratory medicine.

DO GUIDELINES IMPACT CLINICAL PRACTICE AND PATIENT OUTCOMES?

The previously mentioned EU Report also investigated the implementation and impact of guidelines for the management of the most prevalent chronic conditions. They found only two studies that reported effective guideline implementation or impact; five studies showed “partial effectiveness” and three studies did not demonstrate any effectiveness. The BEACH study carried out in Australian general practices investigated pathology test requesting and estimated that 3.1 million tests were reported for Type2 diabetes patients between 2006 and 2008. Seventy two percent of these tests were supported by guideline recommendations, 12.4% were in the grey zone due to unclear guidance and 10.1% were not supported by guidelines (1).

These examples, along with many similar observations published in the literature about guideline implementation, point to the fact that it is not sufficient to develop good evidence-based guidelines and passively disseminate them. Successful translation of the evidence into practice requires a system approach which starts with acknowledging existing gaps in clinical practice and recognising the need for a change, followed by a search for and implementation of

Table 1 Relevance of guideline development standards to laboratory testing-related recommendations

IOM Standard	Explanation	Additional notes relevant to guidelines on laboratory testing
1. Transparency	The guideline development process and its source of funding must be transparent and public	<i>See paper by Kahn et al. in this issue.</i>
2. Conflict of interest	Before guideline panels are established all conflicts of interest must be declared. Chairs and co-chairs should be free from conflicts of interest. Funders of CPGs should not influence the content of the guideline	<i>See paper by Kahn et al. in this issue.</i>
3. Guideline development group composition	Guideline panels should be multidisciplinary involving all key stakeholders targeted by the CPG and methodologists. Patient and consumer involvement should be encouraged.	The involvement of professionals in laboratory medicine should be facilitated in CPG panels where recommendations involve laboratory testing. <i>See paper by Kahn et al. in this issue.</i>
4. Systematic review of the evidence	Guidelines should be based on systematic reviews that meet methodological standards.	Laboratory professionals should be engaged in systemic reviews of diagnostic tests. For recommended tools and checklists see text and the Cochrane DTA and EQUATOR websites.
5. Evidence foundations for and rating the strength of recommendations	Recommendations should have reasoning with clear description of potential benefits and harms and a summary of the evidence behind them. The strength of evidence and the strength of recommendation must be rated. Differences of opinions must be explicitly stated.	The GRADE diagnostic tool is recommended for rating the strength of evidence and the strength of recommendations related to testing. <i>See paper by Don-Wauchope et al. in this issue.</i>

6. Articulation of recommendation	Recommendations must be clear and unambiguous. Strong recommendations should be worded to allow evaluation of compliance.	Recommendations on laboratory testing should consider covering essential items relevant to the correct use and interpretation of laboratory tests (28) <i>See paper by Misra et al. in this issue.</i>
7. External review	External review of draft CPGs should be provided by all relevant key stakeholders, including the public. The guideline panel should address all comments and keep a record on how and why those were incorporated or not in the final recommendations.	<i>See paper by Kahn et al. in this issue.</i>
8. Updating	The CPG publication date, date of systematic evidence review, and proposed time of future update should be documented. The evidence base should be regularly monitored and the CPG updated if significant new evidence emerges that modifies the existing recommendation.	<i>See paper by Kahn et al. in this issue.</i>

Adapted from Institute of Medicine, *Clinical Practice Guidelines We Can Trust*, Standards March 2011 (www.iom.edu/cpgstandards).

a solution through raising awareness and acceptance and leading to adoption and adherence, i.e. the 4A-pipeline of a behavioural change management process.

Active dissemination of guidelines, using leaflets, electronic alerts and advertisements, outreach visits, lectures by respected senior experts and other tools should be coupled with education to raise *awareness* and facilitate *acceptance* of recommendations. However, even acceptance of the evidence does not guarantee that evidence-based recommendations for best practice are adopted and adhered to. Mikan *et al.* have elegantly demonstrated that there is leakage along the awareness–acceptance–adoption–adherence pipeline. Their

study showed that both adoption and adherence were affected by provider and organisational factors. For example, specialists working in large hospitals with better facilities and resources were more likely to adopt and adhere to recommendations than single-handed general practitioners. Laboratories therefore may need to develop different implementation strategies for their hospital and general practitioner clients. It further emphasizes the importance of joint development of laboratory medicine specific clinical recommendations that this study has also found that national or regional recommendations issued by professional organisations were more likely to be accepted and adopted than global or

Table 2 Checklists and tools for laboratory medicine practice guideline developers

Purpose	Tool	Link
Critical appraisal of guidelines	AGREE	www.agreetrust.org/
Grading the strength of evidence and recommendations	GRADE	www.gradeworkinggroup.org
Implementation of guidelines	GLIA	http://nutmeg.med.yale.edu/glia
Standard operating procedure for laboratory medicine practice guideline development	NACB SOP	https://www.aacc.org/~media/files/nacb/nacb_imp_g_sop_jan_2014.pdf?la=en
Guideline development checklist	GRADE-DECIDE	www.guidelinedevelopment.org http://cebgrade.mcmaster.ca/guidecheck.html
Manual for Diagnostic Test Accuracy systematic reviews	Cochrane DTA	http://srdta.cochrane.org/handbook-dta-reviews
Critical appraisal of diagnostic accuracy studies for systematic reviews	QUADAS-2	http://www.bris.ac.uk/media-library/sites/quadas/migrated/documents/quadas2.pdf
Reporting standards for diagnostic accuracy studies	STARD	http://www.equator-network.org
Reporting standards for multivariable prediction model for individual prognosis or diagnosis	TRIPOD	http://www.equator-network.org
Reporting standards for systematic reviews	PRISMA	http://www.equator-network.org

other international guidelines. This study also showed that informed patients can influence adherence to best practice which highlights the importance of guideline implementation strategies that use patient information and empowerment tools. Clear and consistent laboratory

testing-related guidelines, conceived in collaboration with clinical specialists and which are pilot tested and adapted to local settings and equipped with tools and resources for monitoring, achieve higher success with *adoption* and *adherence* (3,21).

The multidimensional and complex nature of guideline implementation strategies is probably best described by a matrix of multifaceted approaches, including 1/ behavioural and educational, 2/ organisational, 3/ policy, and 4/ professional and other incentives and tools that kick the 4A-cycle into action (3). Various evaluations have also concluded that such multifaceted implementation strategies are more likely to succeed than single interventions. Continuous benchmarking of performance, coupled with feedback and education, seem to be the most successful strategies. For more details on guideline implementation and auditing its impact and effectiveness and cost-effectiveness in practice see the papers by Misra et al. and by Collinson, in this issue.

TOWARDS IMPROVED GUIDELINE DEVELOPMENT

The above mentioned shortcomings of guideline development and implementation raise a couple of questions: 1/ Do we need guidelines at all, or should the laboratory profession focus its efforts and resources on producing more high quality research evidence, and probably less low quality guidelines? 2/ Do we need so many guidelines on the same topic, and 3/ is it necessary to have separate guidelines for covering different aspects of care of a clinical condition?

Considering the first question, one may ask, if guidelines are not implemented or applicable to practice and do not have significant impact on health outcomes, why bother developing them? Would it be better to have high quality trials or systematic reviews or evidence summaries in form of well structured, quality rated evidence-tables that would provide a universal answer to clinically important questions? This might be particularly relevant in laboratory medicine, where systematic reviews,

conducted solely for the purposes of guidelines or economic analyses, are often of poorer quality than single overviews performed by experts trained in evidence-based medicine (22,23). Kahn and Gale also argue whether we need so many guidelines on the same topic and whether we should move away from guidelines that are too generic and directed toward patient populations and replace those by computer generated, individualised guidelines where the evidence is provided as a backbone for local discussions and formulation of local policies on best clinical practice (24).

So, there is an increasing argument for simply providing better evidence and evidence reviews. However, for doing so, laboratory professionals should obtain more skills in systematically reviewing the diagnostic literature which itself has a number of methodological challenges. Various manuals and tools assist in writing systematic reviews related to diagnostic testing (Table 2). For example, a Cochrane Working Group issued a comprehensive handbook for systematic reviews of diagnostic test accuracy (<http://srdta.cochrane.org/handbook-dta-reviews>). The so-called QUADAS tool is a very useful resource for appraising diagnostic accuracy studies for systematic reviews (25). Several reporting standards, such as STARD for diagnostic accuracy studies, TRIPOD for a multivariable prediction model for individual prognosis or diagnosis, and PRISMA for systematic reviews in general can be found on the Equator Network's website (<http://www.equator-network.org>). The Institute of Medicine has also issued methodological standards for producing high quality systematic reviews (http://www.nap.edu/catalog.php?record_id=13059).

Regarding the second question, undoubtedly we have far too many guidelines, often covering the same topic. At the time of writing this article, 2417 CPGs are available in the Agency

for Healthcare Research and Quality's (AHRQ) National Guideline Clearinghouse website (<http://www.guideline.gov/index.aspx>) and 84 CPGs are under development. On diabetes mellitus alone, there are 454 hits for CPGs in the same database. For this phenomenon of multiple guidelines, Kahn and Gale offer some explanations and a solution. If a new guideline is developed for a topic that is already covered by a guideline elsewhere, the organisation should provide a rationale why a new guideline is needed; simply approve the existing guideline if recommendations are the same; or explain how and why the new guideline differs from the previous one (24). The AHRQ also offers guideline synthesis reports that compare the scope, content and the corresponding strength of evidence of various CPGs on the same topic (<http://www.guideline.gov/compare/index.aspx>).

With increasing rigour for development CPGs are becoming too complex and too long. There is an ongoing discussion whether testing-related recommendations should be developed by subspecialty societies, such as AACC's NACB, or whether recommendations on testing should be part of CPGs and developed jointly with clinical societies. On the one hand, guidelines produced by specialty societies are reported to be of lower methodological quality compared to those produced by major guideline organisations that have well-defined processes, rigorous methodologies and adequate resources to hire expertise for evidence-based guideline development (26,27). On the other hand, laboratory testing-related information is not easy to locate and pre- and post-analytical information, important for the appropriate requesting and use of tests, is rarely provided in CPGs. Inappropriate coverage of laboratory testing related information in CPGs has been shown by the Guideline Working Group of the European Federation of Clinical Chemistry and

Laboratory Medicine. For more comprehensive coverage of laboratory-related items in CPGs, this group has suggested a detailed checklist of 33 preanalytical, 37 analytical and 10 post-analytical items and they also provided a reduced list of minimum requirements (28). For more details see the paper by Misra et al., in this issue.

CONCLUSIONS

Whilst clinical practice guidelines aim to close the gap between research and practice, the appearance of so many guidelines seems to have created a new gap between their development and utility in practice. Poor quality and lack of explicitness of recommendations on laboratory testing call for methodological and reporting standards for guidelines. A transparent and explicit evidence-grading scheme and international collaboration of guideline development activities are needed to increase the validity, applicability and cost-effectiveness of recommendations related to the use of laboratory tests in clinical practice.

ACKNOWLEDGEMENTS

Aspects of this review have been discussed in and influenced by the work of the Test Evaluation Working Group and the Guidelines Working Group of the European Federation of Clinical Chemistry and Laboratory Medicine. Roche Diagnostics is thanked for supporting the Test Evaluation Working Group with an independent educational grant to the European Federation of Clinical Chemistry and Laboratory Medicine.

REFERENCES

1. Bayram C, Britt H, Miller G, Valenti L 2009. Evidence-practice gap in GP pathology test ordering: a comparison of BEACH pathology data and recommended testing. Family Medicine Research Centre School of Public Health.

- The University of Sydney, June 2009;pp205. [https://www.health.gov.au/internet/main/publishing.nsf/Content/9C300FE48F876E95CA257BF0001ACE0E/\\$File/Evidence-practice%20gap%20in%20GP%20pathology%20test%20ordering.pdf](https://www.health.gov.au/internet/main/publishing.nsf/Content/9C300FE48F876E95CA257BF0001ACE0E/$File/Evidence-practice%20gap%20in%20GP%20pathology%20test%20ordering.pdf) (Last accessed 2 June 2015)
2. Field MJ, Lohr KN. (eds) Guidelines for Clinical Practice: from development to use, Institute of Medicine, National Academy Press, Washington, DC.1992
3. Horvath AR. From Evidence to Best Practice in Laboratory Medicine. Clin Biochem Rev 2013;34:47-60.
4. Evidence-Based Health Care. How to make health policy and management decisions. Ed.: JAM Gray, Churchill Livingstone, 1997;pp270
5. von Karsa L, Patnick J, Segnan N, Atkin W, Halloran S, et al. European guidelines for quality assurance in CRC screening and diagnosis Overview and introduction to the full Supplement publication. Endoscopy 2013;45:51-59.
6. Klabunde CN, Joseph DA, MD, King JB, White A, Plescia M. Vital Signs: Colorectal Cancer Screening Test Use — United States, 2012. Morbidity and Mortality Weekly Report 2013; 62(44):881-888. <http://www.cdc.gov/mmwr> (Last accessed 2 June 2015)
7. von Karsa L, Anttila A, Ronco G, Ponti A et al. Cancer Screening in the European Union Report on the implementation of the Council Recommendation on cancer screening, First Report, European Commission 2008; pp160.
8. McDowell SE, Thomas SK, Coleman JJ, Aronson JK, Ferner RE. A practical guide to monitoring for adverse drug reactions during antihypertensive drug therapy. J R Soc Med 2013;106:87–95.
9. Kim EH, Andriole GL. Prostate-specific antigen-based screening: controversy and guidelines. BMC Medicine 2015;13:61. <http://www.biomedcentral.com/1741-7015/13/61> (Last accessed 2 June 2015)
10. Mpondo BCT, Ernest A, Dee HE. Gestational diabetes mellitus: challenges in diagnosis and management. J Diabetes Metab Disord. 2015;14:42. <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4430906> (Last accessed 2 June 2015)
11. Nankervis A, Conn J. Gestational diabetes mellitus - Negotiating the confusion. Australian Family Physician 2013;42(8):528-531. <http://www.racgp.org.au/afp/2013/august/gestational-diabetes-mellitus/> (Last accessed 2 June 2015)
12. Watine J, Friedberg B, Nagy E, Onody R, Oosterhuis W, Bunting P, Charet J-C, Horvath AR. Conflict between guideline methodologic quality and recommendation validity: a potential problem for practitioners. Clin Chem 2006;1(52): 65-72.
13. Nagy E, Watine J, Bunting P, Onody R, Oosterhuis W, Rogic D, Sandberg S, Boda K, Horvath AR. Do guidelines for the diagnosis and monitoring of diabetes mellitus fulfill the criteria of evidence-based guideline development? Clin Chem 2008;54(11):1872-82.
14. Alonso-Coello P, Irfan A, Solà I, Gich I, Delgado-Noguera M, Rigau D, Tort S, Bonfill X, Burgers J, Schunemann H. The quality of clinical practice guidelines over the last two decades: a systematic review of guideline appraisal studies. Quality and Safety in Health Care, 2010;19(6):e58.
15. Don-Wauchope AC, Sievenpiper JL, Hill SA, Iorio A. Applicability of the AGREE II instrument in evaluating the development process and quality of current National Academy of Clinical Biochemistry guidelines. Clin Chem. 2012;58(10):1426-37.
16. Gopalakrishna G, Langendam MW, Scholten RJPM, Bossuyt PMM, Leeflang MMG. Guidelines for guideline developers: a systematic review of grading systems for medical tests. Implementation Science 2013;8:78 <http://www.implementationscience.com/content/8/1/78> (Last accessed 2 June 2015)
17. Brouwers M, Kho ME, Browman GP, Cluzeau F, Feder G, Fervers B, Hanna S, Makarski J on behalf of the AGREE Next Steps Consortium. AGREE II: Advancing guideline development, reporting and evaluation in healthcare. Can Med Assoc J. 2010;182:E839-842. The AGREE II Instrument is accessible at: <http://www.agreetrust.org/> (Last accessed 2 June 2015)
18. Legido-Quigley H, Panteli D, Car J, McKee M, Busse R. Clinical guidelines for chronic conditions in the European Union. Observatory Studies Series 30. The European Observatory on Health Systems and Policies. 2013;pp256.http://www.euro.who.int/_data/assets/pdf_file/0009/195876/Clinical-Guidelines-for-Chronic-Conditions-in-the-European-Union.pdf (Last accessed 2 June 2015)
19. Schünemann HJ, Wiercioch W, Etzeandía I, Falavigna M, Santesso N, Mustafa R, Ventresca M et al. Guidelines 2.0: systematic development of a comprehensive checklist for a successful guideline enterprise. CMAJ 2014;186(3):E123-142.
20. Institute of Medicine, Clinical Practice Guidelines We Can Trust. Accessible at: www.iom.edu/cpgstandards (Last accessed 2 June 2015)
21. Mikan S, Burls A, Glasziou P. Patterns of 'leakage' in the utilisation of clinical guidelines: a systematic review. Postgrad Med J 2011;87:670-9.
22. Vigna-Taglianti F, Vineis P, Liberati A, Faggiano F. Quality of systematic reviews used in guidelines for oncology practice Ann Oncol 2006;17:691-701,

23. Mallett S, Deeks JJ, Halligan S, Hopewell S, Cornelius V, Altman DG. Systematic reviews of diagnostic tests in cancer: review of methods and reporting. *BMJ* 2006;333:413-19.
24. Kahn R, Gale EAM, Gridlocked guidelines for diabetes. *Lancet* 2010;375:2203-4.
25. Whiting PF, Rutjes AWS, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, Leeflang MMG, Sterne JAC, Bossuyt PMM, the QUADAS-2 Group. QUADAS-2: A Revised Tool for the Quality Assessment of Diagnostic Accuracy Studies. *Ann Intern Med.* 2011;155:529-36.
26. Burgers JS, Cluzeau FA, Hanna SE, Hunt C, Grol R: Characteristics of high quality guidelines: evaluation of 86 clinical guidelines developed in ten European countries and Canada. *Int J Technol Assess Healthcare* 2003;19(1):148–157.
27. Grilli R, Magrini N, Penna A, Mura G, Liberati A: Practice guidelines developed by specialty societies: the need for a critical appraisal. *Lancet* 2000;355:103–105.
28. Aakre KM, Langlois MR, Watine J, Barth JH, Baum H, Collinson P, Laitinen P, Oosterhuis WP. Critical review of laboratory investigations in clinical practice guidelines: proposals for the description of investigations. *Clin Chem Lab Med* 2013;51:1217-26.