Guideline Quality and Guideline Content: Are They Related?

Evidence-based medicine (EBM) challenges clinicians and laboratory professionals to make rational decisions in healthcare. Clinical practice guidelines embody the principles of EBM in making options and choices explicit, considering the scientific evidence and resources available (1). The evidence is often limited or controversial, however, and resources differ among countries and regions. Therefore, the translation of research evidence into recommendations for clinical practice is not straightforward. Judgment beyond the evidence is necessary, taking into consideration the balance between benefits and harms and risks, patients’ views and preferences, and potential organizational and financial barriers (2). Addressing these issues requires careful discussions within a working group that includes representatives from all relevant disciplines. In contrast to the development of systematic reviews, guidelines cannot be produced in an academic “ivory tower” by a few experts.

When the evidence is not strong or not fully applicable to the patient population targeted in the guideline, the working group may depart from the evidence. In laboratory medicine, high-level diagnostic evidence is particularly scarce, and the link between diagnostic tests and the patient population targeted in the guideline, the academic “ivory tower” by a few experts.

Evidence-based medicine (EBM) challenges clinicians and laboratory professionals to make rational decisions in healthcare. Clinical practice guidelines embody the principles of EBM in making options and choices explicit, considering the scientific evidence and resources available (1). The evidence is often limited or controversial, however, and resources differ among countries and regions. Therefore, the translation of research evidence into recommendations for clinical practice is not straightforward. Judgment beyond the evidence is necessary, taking into consideration the balance between benefits and harms and risks, patients’ views and preferences, and potential organizational and financial barriers (2). Addressing these issues requires careful discussions within a working group that includes representatives from all relevant disciplines. In contrast to the development of systematic reviews, guidelines cannot be produced in an academic “ivory tower” by a few experts.

When the evidence is not strong or not fully applicable to the patient population targeted in the guideline, the working group may depart from the evidence. In laboratory medicine, high-level diagnostic evidence is particularly scarce, and the link between diagnostic tests and the patient population targeted in the guideline, the academic “ivory tower” by a few experts.

High-quality guidelines are based on evidence as well as a broad consensus of opinions, which facilitates the acceptance and effective use of the guideline in the target group (6). To ensure high quality, guidelines should be developed within a structured and coordinated program according to the principles of evidence-based guideline development (7). Explicit reporting of the methods and procedures followed will improve the guideline quality (8). On the other hand, lack of information about the methods and procedures does not automatically mean that the recommendations in the guideline are not valid.

The AGREE (Appraisal of Guidelines Research and Evaluation) Instrument, first published in 2001 (9), is a tool for assessing the quality of clinical guidelines according to 23 criteria, grouped into 6 domains (scope and purpose, stakeholder involvement, methodology, clarity and presentation, applicability, and editorial independence). The instrument is available in more than 10 languages and is currently used in many countries around the world. The AGREE criteria mainly concern the methods used for developing the guideline and the quality of the reporting. The clinical content of the recommendations and the quality of the supporting evidence are not addressed, however (10), a deficit common in appraisal tools for clinical guidelines (11). One might assume that a guideline with a high quality score would contain valid recommendations that are not in conflict with the best available evidence. If this were true, quality assessment of the evidence behind the recommendations would not be needed, and much appraiser time would be saved.

In this issue of Clinical Chemistry, Watine et al. (12) report the results of their test of the correlation between guideline quality and clinical validity of recommendations for 11 guidelines on a specific topic (non-small cell lung cancer) in laboratory medicine. Guideline quality was measured with the AGREE Instrument summarized in a global score. The validity of the recommendations was assessed by comparison with the results of a systematic review performed by the authors themselves. This study is the first to test the relationship between the AGREE quality scores and the clinical content of guidelines, a process that could be considered as a next step in validating the AGREE Instrument. The results revealed that there was no relationship: “good” as well as “not so good” guidelines contained “good” recommendations, corresponding with the evidence of the systematic review, and even “good” guidelines included “not so good” recommendations.

These findings confirm the variability of the translation of evidence into practice recommendations by different guideline groups. Such variability, however, is not necessarily undesirable. Limited availability of resources could explain the fact that some guidelines do not recommend the whole series of laboratory tests determined to be useful according to the evidence from the systematic review. On the other hand, a lack of budget restrictions might account in part for the recommended use of tumor markers, whereas this was not recommended in the systematic review. Ideally, these considerations should have been made explicit in the guideline. Guideline documents offer the opportunity to include balanced thoughts behind the recommendations, to identify gaps in knowledge, and to include different options when the evidence is unclear or lacking. Such information is a better alternative to nonspecific recommendations that could potentially harm patients (13).

One important limitation of the study by Watine et al. (12) is that they used their own systematic review as a gold standard. This would be appropriate if sufficient evidence were available, including data from high-quality trials or prospective cohorts in the case of specific diagnostic research questions. Most of their conclusions, however, were not based on the highest level of evidence but on validation by experts. Thus, their review itself could be considered a guideline based on evidence and consensus, inevitably including some subjective judgments. Moreover, guidelines are particularly needed in areas of uncertainty. If the evidence is clear, guidelines would not have an added value compared with a review or textbook.

Some variation in guidelines is acceptable because of context-specific decisions. One should therefore be cautious in promoting international guidelines (14). Nevertheless, international standards similar to those for randomized clinical trials and diagnostic studies could be
developed for guideline quality and reporting (15, 16). The Guidelines International Network, founded in 2002, provides good opportunities to exchange methods and new approaches to guideline development (17). Further validation of the AGREE Instrument would be the next step to achieve more international consensus about guideline quality and methodology.

References

Jako S. Burgers
Centre for Quality of Care Research
Radboud University Nijmegen Medical Centre
Nijmegen, The Netherlands
and
Dutch Institute for Healthcare Quality CBO
Utrecht, The Netherlands

Address for correspondence: Dutch Institute for Healthcare Quality CBO, Churchilllaan 11, PO Box 2064, 3502 LB Utrecht, The Netherlands. Fax 31-30-294-3644; e-mail j.burgers@cbo.nl.

DOI: 10.1373/clinchem.2005.059345