MAYO CLINIC

REVIEW

Clinical Practice Guidelines: A Primer on Development and Dissemination

M. Hassan Murad, MD, MPH

Abstract

Trustworthy clinical practice guidelines should be based on a systematic review of the literature, provide ratings of the quality of evidence and the strength of recommendations, consider patient values, and be developed by a multidisciplinary panel of experts. The quality of evidence reflects our certainty that the evidence warrants a particular action. Transforming evidence into a decision requires consideration of the quality of evidence, balance of benefits and harms, patients' values, available resources, feasibility of the intervention, acceptability by stakeholders, and effect on health equity. Empirical evidence shows that adherence to guidelines improves patient outcomes; however, adherence to guidelines is variable. Therefore, guidelines require active dissemination and innovative implementation strategies.

 ■ Mayo Clin Proc. 2017;92(3):423-433

Clinical practice guidelines are systematically developed statements that intend to assist clinicians and patients in making decisions about appropriate health care in specific circumstances.[1] Guidelines aim to improve the quality of patient care by encouraging interventions of proven benefit and discouraging the use of ineffective or potentially harmful interventions; to reduce unnecessary variation in practice; to lessen disparities; to empower patients; and to influence public policy.[2]

Production of guidelines has skyrocketed during the past 30 years. Currently, the library of the Guidelines International Network has 6187 documents from 76 countries, and the National Guideline Clearinghouse in the United States has 2017 guideline summaries.[3,4] Guidelines are critical for developing disease performance measures and defining high-value care.

This primer includes a description of the modern approach to developing guidelines; the criteria for trustworthy guidelines; advice on how clinicians can appraise, interpret, and implement practice recommendations; challenges and limitations of guidelines; and a future research agenda to address current knowledge gaps.

HISTORICAL PERSPECTIVE

Until the 1970s, medical actions were indirectly regulated through the training and credentials granted by medical schools or state authorities; however, such credentialing proved to be an insufficient guarantee of quality.[5] Further standardization and organization of the medical profession necessitated the development of guidelines. Guidelines in their current form started in the 1970s and were primarily based on the consensus of expert panels (eg, the National Institutes of Health Consensus Development Program).[6] Experts recommended management approaches they have used in their practice and cited references they recalled or were able to identify without an explicit systematic search. With the emergence of evidence-based medicine as a principle for decision making in the 1980s and coining of the term in 1991,[7] more rigorous approaches for guideline development have emerged. The next generation of guidelines would emphasize research evidence over opinion and base recommendations on the design of studies contributing evidence on benefits and harms of interventions. For example, the American College of Cardiology/American Heart Association used ratings of A, B, and C that were exclusively dependent on study design (level A, multiple randomized trials or meta-analyses; level B, a single trial or nonrandomized studies; and level C, consensus opinion of experts, case studies, or standard of care).[8] In 2000, many different guideline systems existed, which

For editorial comment, see page 327

From the Division of Preventive Medicine, Mayo Clinic, Rochester, MN.

ARTICLE HIGHLIGHTS

- Trustworthy clinical practice guidelines require a systematic review to select the best available evidence and should explicitly evaluate the quality of evidence.
- Factors that reduce the quality of evidence are risk of bias, indirectness, inconsistency, imprecision, and likelihood of publication and reporting bias.
- Transforming evidence into a decision requires consideration of the quality of evidence, balance of benefits and harms, patients' values, resources, feasibility, acceptability, and equity.
- Empirical evidence shows that guidelines improve patient outcomes; however, guidelines require active dissemination and innovative implementation strategies.

was confusing for stakeholders. The 6 most prominent systems had low reproducibility of judgments and did not fit the needs of all stakeholders.[9]

Subsequently, it became apparent that study design (a surrogate for the risk of bias) was insufficient.[9] Studies with the same design (eg, randomized trials) can have high risk of bias or low risk of bias. Factors other than risk of bias affect certainty in the evidence (eg, precision of estimates of effect, consistency of effect across studies). Moreover, factors other than evidence (eg, patient values) affect decision making.[10]

Considering these challenges and to unite many of these frameworks, the GRADE approach (Grading of Recommendations, Assessment, Development, and Evaluation) was developed in 2003. It advanced guideline methodology further by providing a framework for rating the quality of evidence based on 8 distinct domains (as opposed to study design only). The construct of the quality of evidence was then defined to reflect the extent of our confidence that the estimate of an effect is adequate to support a particular decision or recommendation.[11] An evidence to decision framework based on 8 criteria (as opposed to intuitive or global judgment) was provided by GRADE to assist in developing actions based on evidence.[12] Empirical evaluation showed that recommendations with a rating of A (based on multiple randomized trials or meta-analyses) would have been rated using GRADE as high, moderate, low, and very low.[13] Thus, this newer approach uncovered additional factors affecting the quality of evidence that were otherwise implicit. Judgments made using GRADE had good reproducibility and reliability compared with intuitive or global judgments and were consistent among raters (interrater reliability of 0.72)[14] even when panel members had short training in GRADE (two 1-hour didactic sessions).[15] GRADE was adopted by more than 100 organizations and has become, to some extent, the gold standard (when other systems are used,[16,17] they usually depend on GRADE domains and components).

In 2011, the National Academy of Medicine (formerly the Institute of Medicine) published criteria for trustworthy guidelines that greatly overlapped with GRADE (in emphasis on the systematic review process, rating the quality of evidence and strength of recommendations, and considering nonevidence factors).[18] These criteria were highly disseminated and cited, and they motivated guideline developers to improve the rigor of guidelines (Table 1). Similar criteria were also produced by the Guidelines International Network,[19] the National Institute for Health and Clinical Excellence in the United Kingdom,[20] and the World Health Organization.[21]

EVALUATING THE QUALITY OF EVIDENCE

A systematic review is a mechanism to reduce the risk of biased selection of evidence and should be conducted once the scope and preliminary questions of the guideline are determined. Meta-analysis may or may not be appropriate, but a systematic review is always needed. In the context of a guideline, the quality of evidence (also called certainty in the evidence, strength of the evidence, and confidence in the effect estimates) reflects the extent of our confidence that the estimates of an effect are adequate to support a particular decision or recommendation.[11] A good starting place to determine this confidence depends on study design (higher confidence in randomized trials and lower confidence in observational studies). There are, however, other factors that can moderate this initial rating of confidence. We are less confident of the effect of an intervention on benefits and harms (1) if the evidence is derived from studies with methodological limitations

Mayo Clin Proc. ■ March 2017;92(3):423-433 ■ http://dx.doi.org/10.1016/j.mayocp.2017.01.001
www.mayoclinicproceedings.org