



## Evidence-Based Guidelines—An Introduction

Wendy Lim,<sup>1</sup> Donald M. Arnold,<sup>2</sup> Veronika Bachanova,<sup>3</sup> Richard L. Haspel,<sup>4</sup> Rachel P. Rosovsky,<sup>5</sup> Andrei R. Shustov,<sup>6</sup> and Mark A. Crowther<sup>7</sup>

<sup>1</sup>McMaster University, Hamilton, Ontario, Canada; <sup>2</sup>Michael G. DeGroot School of Medicine, McMaster University, Hamilton, Ontario, Canada; <sup>3</sup>Department of Hematology, Oncology and Transplantation, University of Minnesota, Minneapolis, MN; <sup>4</sup>Beth Israel Deaconess Hospital, Department of Pathology, Boston, MA; <sup>5</sup>Massachusetts General Hospital and Harvard Medical School, Boston, MA; <sup>6</sup>University of Washington, Division of Hematology and Fred Hutchinson Cancer Research Center, Clinical Research Division, Seattle, WA; <sup>7</sup>St. Joseph's Hospital and McMaster University, Hamilton, Ontario, Canada

**Recommendations in the form of clinical practice guidelines are increasingly common. Clinical guidelines are systematically developed statements designed to help administrators, practitioners and patients make decisions about appropriate health care for specific circumstances. In North America, guidelines developed by professional societies, government panels and cooperative groups are frequently used to**

**measure quality, to allocate resources and to determine how health care dollars are spent. For clinicians, guidelines provide a summary of the relevant medical literature and offer assistance in deciding which diagnostic tests to order, which treatments to use for specific conditions, when to discharge patients from the hospital, and many other aspects of clinical practice.**

Applying evidence to clinical practice is a requirement for best patient care. “Evidence-based medicine” takes many forms and is frequently viewed with some skepticism by practicing clinicians. However, fundamental to evidence-based practice is the concept that, although evidence can recommend particular diagnostic strategies, treatments or management plans, each plan must be individualized to reflect the specific characteristics of individual clinical circumstances. That having been said, government and third-party payers are increasingly using evidence summaries to guide reimbursement decisions—interventions with good quality evidence are thus more likely to be supported by payers than interventions lacking such evidence. Fundamental to understanding evidence-based practice and guideline development are systematic reviews. Systematic reviews use well-defined and reproducible literature search strategies to identify evidence that informs clinical problems; the data is then assessed for its methodological rigor and, when of sufficient quality and quantity, it may be confined mathematically using meta-analysis techniques, which provide more accurate estimates of effects as they incorporate larger numbers of patients than the source studies.

Evidence-based guidelines use systematic reviews to inform specific clinical circumstances. Clinical questions are defined and systematic literature review is performed to identify evidence that addresses those clinical questions. The evidence is then summarized, graded (an indication of its strength and methodological rigor) and presented. Evidence-based guidelines can improve the quality of care of patients by supporting interventions of proven benefit,

while discouraging interventions that are ineffective or potentially harmful. However, guidelines can also be misleading and cause harm, particularly if recommendations are based on an incomplete or flawed dataset, if they are biased, or if the development process is incorrectly carried out. Rigorously developed evidence-based guidelines aim to minimize these potential harms.

The science and methodology of guideline development has evolved, with a focus on providing recommendations based on evidence from the medical literature. Except in clinical settings where there is limited evidence, guidelines based on consensus or expert opinion alone are becoming less acceptable. Evidence-based guidelines use the principles of evidence-based medicine: conducting a comprehensive literature search, critically appraising the quality of the evidence and generating recommendations while considering patients’ preferences and values. Additionally, an increasing number of evidence-based guidelines grade the quality of the evidence that the recommendations are based upon. Grading medical evidence provides users of the guidelines with a more uniform interpretation of the recommendations and describes when recommendations are clearly beneficial or harmful, or where the benefits and risks are uncertain. The grading system also clearly indicates the level of evidence (study design) supporting the recommendations. Transparent reporting of the process of guideline development, carefully conducted systematic searches with assessment of study quality, and use of a grading system are components of high-quality evidence-based guidelines, which aim to provide users of the guide-

lines with a useful tool to improve patient quality of care.

A frequent misconception is that guidelines and “evidence-based practice” is limited to clinical areas with a plethora of large Phase III studies; in fact, rigorous guidelines can be formulated for any clinical area, irrespective of the “quality” of the underlying data. Rare diseases or infrequently used treatments that receive lower grade recommendations (for example, a 2C recommendation using the Grade system) may be as clinically useful as well accepted interventions for common diseases that have been well studied and receive a 1A recommendation. The lower “grade” recommendation simply reflects the lack of good quality supporting data and should not be misinterpreted as a recommendation against something.

This chapter provides a brief introduction to Evidence Based Guidelines and is followed by 5 brief evidence-based summaries<sup>1-5</sup> written by ASH Clinical Research Training Institute graduates or Junior ASH faculty members in concert with the author of the corresponding section of *Hematology 2008*, the ASH Education Book.

## Background

Clinical guidelines are systematically developed statements designed to help practitioners and patients make decisions about appropriate health care for specific circumstances.<sup>6</sup> Although designed to improve the quality of care they can also serve as a useful tool for practitioners faced with mounting evidence that is often presented in obscure journals in difficult to digest formats.<sup>7</sup>

Clinical guidelines must summarize the evidence available to guide informed decisions. This is challenging since many clinical areas lack high quality evidence and yet guidelines have been formulated—guidelines based on consensus or a partial review of the literature are more prone to bias than systematic guidelines. Overcoming this limitation requires a formalized description of the methodology of the review and the quality of the evidence upon which it is based. This has led to increasing interest in evidence-based guidelines, which critically appraise the available supporting evidence and highlight its strengths and weaknesses. The systematic examination of evidence promoted by evidence-based medicine is done through a comprehensive search of the literature, critical appraisal of the quality of the evidence and interpretation of the findings in light of patients’ preferences and societal values.<sup>8</sup> This process reduces the emphasis on unsystematic clinical experience and pathophysiological rationale. The principles of evidence-based medicine have become an integral part of undergraduate, postgraduate and continuing medical education and its use has been shown to have a positive effect on practice and patient care.<sup>9</sup>

Evidence-based guidelines apply the principles of evidence-based medicine to the process of guideline development. Generation of the recommendations involves

both content and methodology experts, and the process is clearly defined and reproducible. The first step in the process of evidence-based guideline development is defining the clinical question that the guideline will address. This is followed by defining the eligibility criteria for the studies that will be included in the guideline recommendations. A systematic search of the literature is then conducted and the evidence is evaluated. In developing recommendations, the likely benefits, risks, inconvenience and costs associated with each treatment must be considered in addition to addressing patients’ underlying values and preferences.<sup>10</sup> The quality of the data supporting the recommendations is evaluated and is reflected in a grading system that describes the strength of the recommendation and the quality of the supporting evidence. This process ultimately results in the systematic development of recommendations that incorporate evidence with patients’ preferences and values and indicates the quality of the evidence.

An example of an evidence-based guideline is the publication by the American College of Chest Physicians (ACCP) Conference on Antithrombotic and Thrombolytic Therapy.<sup>11</sup> The process of generating recommendations in this document will be used to highlight the methodology of evidence-based guideline development.

## Defining the Clinical Question

Evidence-based guidelines are designed to summarize the evidence and address a specific question regarding a medical condition. Defining the question is a critical first step, and involves clearly defining the patient population, intervention, comparison and outcome. This in turn will define the eligibility criteria that will be used in the systematic search. Defining and refining the clinical question requires careful consideration of patient-specific and disease-specific factors, as well as treatment characteristics and the impact of these on the outcomes of interest. Patient population, the intervention, comparator and outcome must be defined in turn. In defining a drug intervention, the route and frequency of administration, and the use of additional treatments (co-interventions) are important considerations. Alternative treatments or management strategies may be included. The intervention may be compared to no treatment, placebo, or an alternate treatment to estimate the treatment effect. The outcome is usually based on clinically important endpoints since one of the cornerstones of evidence-based medicine is that the evidence is relevant to individual patients.

Patient- and disease-specific factors may include the extent or stage of disease, any prior treatments received, risk factors or prognostic factors and comorbidities. For example, a guideline addressing the treatment of patients with multiple myeloma will have different recommendations in patients presenting with newly diagnosed multiple myeloma compared to patients with relapsed disease,

and recommendations may differ in patients who are younger and are candidates for hematopoietic stem cell transplantation, an option that carries higher risk among patients who are older or have multiple comorbidities. Thus, in this example, a clear definition of the target population (patients with newly diagnosed/untreated multiple myeloma versus patients with relapsed multiple myeloma) and the treatment (comparing different chemotherapeutic regimens or comparing transplantation with chemotherapy) must be provided so the guidelines can be utilized for the appropriate patient population. Further refinements in defining the target population by patient or treatment characteristics will result in more specific recommendations, at the cost of decreasing the generalizability of the guidelines.

### **Systematic Searching of the Literature**

As described above, a carefully defined clinical question will provide the framework for defining eligibility criteria for the literature search. The search is usually conducted by trained librarians and methodology experts using a comprehensive search strategy. In addition to having eligibility criteria for the population, interventions and outcomes, some guidelines may additionally limit study inclusion by the type of study design. For example, authors who are evaluating treatment studies may consider only data from randomized controlled trials. This is because the design of a randomized trial minimizes bias, which can influence study results; thus, a well performed randomized controlled trial (RCT) produces more accurate treatment effect estimates. Observed treatment effects are usually larger in non-randomized designs. However, for many clinical questions, randomized trials may not exist or may not be the most appropriate design to use for addressing a specific clinical question. In these cases, observational studies will be evaluated for inclusion.

Systematic searches are typically performed using a broad search strategy. Limitations may be applied to the search, such as excluding studies published in languages other than English or excluding conference abstracts. The process used for the literature search is reported in a transparent manner and is thus reproducible. The databases used for the search are specified, and may include MEDLINE, Embase, the Cochrane Library (Cochrane Database of Systematic Reviews, the Database of Abstracts of Reviews of Effects, Cochrane Central Register of Controlled Trials) and other databases. International groups such as the Cochrane Collaboration (<http://www.cochrane.org>, accessed May 4 2008), the York Centre for Reviews and Dissemination ([www.york.ac.uk/inst/crd/](http://www.york.ac.uk/inst/crd/), accessed May 4 2008) and the Cancer Care Ontario Program in Evidence-Based Care (<http://www.cancercare.on.ca/english/toolbox/qualityguidelines>, accessed May 4 2008) maintain databases of systematic reviews that can provide guideline authors and clinicians with comprehensive overviews of the literature.

### **Critical Appraisal of Study Quality and Grading the Medical Literature**

Systematic reviews are an important component to developing evidence-based guidelines but do not provide sufficient information for making informed decisions in health care. This is because clinicians and others who utilize guidelines draw their own conclusions about the strength of the recommendations and their applicability to individual patients. Clinical guidelines should therefore assess the methodologic quality of the studies that form the basis of their guideline recommendations to standardize judgments about how convincing the evidence is. This process highlights the gaps in the literature and identifies where further good quality research is required.

Systems have evolved to grade the quality of the evidence and the strength of guideline recommendations. Systematic grading can minimize bias and help users of the guideline interpret the treatment recommendations in a more consistent manner. Several different grading systems exist, although an international collaboration known as the Grading of Recommendations Assessment, Development and Evaluation (GRADE) Working Group is working to establish a common system for grading medical evidence.<sup>12</sup> The Scottish Intercollegiate Guideline Network (SIGN) is another multidisciplinary group that develops evidence-based guidelines for the National Health Service (NHS) in Scotland and has proposed a grading system for evaluating evidence.<sup>13</sup>

In the Eighth ACCP Conference on Antithrombotic and Thrombolytic Therapy, recommendations are divided into *strong recommendations*, where clinicians are very certain that the benefits do, or do not, outweigh risks and burdens, or *weak recommendations*, where there is uncertainty as to the magnitude of the benefits and risks. Strong recommendations are given a Grade 1, and weak recommendations are given a Grade 2. Recommendations are then further classified into the quality of the evidence supporting the recommendation. Consistent results from randomized trials result or observational studies with very strong effects result in grade A recommendations, poorer quality results from randomized trials result in grade B recommendations, and weaker studies result in grade C recommendations.

### **Summarizing the Evidence into Guideline Recommendations**

The best available evidence is used to formulate guideline recommendations. Evidence is considered high quality if the study design minimizes potential biases. Study designs can be considered in a hierarchy, with rigorously performed systematic reviews and meta-analyses of randomized controlled trials considered the “highest quality” evidence. In ascending order the likelihood of bias influencing results increases with randomized controlled trials, non-randomized

interventional studies, cohort studies, case-controlled studies, case series and case reports and, finally, expert opinion. This hierarchy can be affected by the quality of the contributing studies; thus, a well done observational study demonstrating a very strong treatment effect may provide better evidence than an underpowered or poorly performed RCT that failed to find a significant treatment effect.

In presenting an evidence-based review, the evidence is summarized and patients' values and preferences are then factored into the recommendation. Values and preferences are subjective and individualized, and may vary among patients and the guideline authors. Furthermore, patients' values and preferences are usually unavailable for most clinical situations outlined in the guidelines. To address this limitation, the ACCP Conference on Antithrombotic and Thrombolytic Therapy guidelines had the recommendations reviewed by a diverse panel, including research methodologists, practicing generalists and specialists. The recommendations that were subject to values and preferences were identified and underlying values and preferences were specified in the guidelines to the greatest extent possible.

### **Advantages of Evidence-Based Guidelines for Clinical Practice**

Clinical practice guidelines have advantages for patients, health care providers and medical researchers. For patients, guidelines have been shown to improve quality of care.<sup>14</sup> Guidelines that support the use of interventions that are clearly of benefit and discourage ineffective or harmful interventions may reduce morbidity and mortality and improve quality of life. Guidelines can also increase the consistency of care, such that patients with identical clinical problems receive consistent and appropriate treatment, irrespective of where and by whom they are treated. Lastly, guidelines can benefit patients by influencing public policy, particularly by highlighting areas where research is lacking and pointing out under-recognized health problems and clinical services. In some settings support for a practice or intervention may lead to its funding by governments or health insurance providers.

Health care providers can also benefit from guidelines since they can improve the quality of clinical decisions. Guidelines offer recommendations for clinicians who may be uncertain how to best diagnose or treat a specific condition, and can educate clinicians by providing a summary of the supporting evidence behind the recommendations. Guidelines can also be used as a reference for quality assurance projects and initiatives, such as preprinted standardized orders and care pathways. They can be used for practice audits, where clinicians' or hospitals' practices can be compared to a common standard of care, to demonstrate where improvements can be made.

Researchers also benefit from evidence-based guidelines that demonstrate where knowledge gaps exist and where further research is required. Funding agencies may use guidelines to identify areas of research in need of support.

### **Limitations of Evidence-Based Guidelines**

Guidelines also have limitations.<sup>7</sup> Clearly, if the recommendation is inaccurate it can increase harm to patients or result in suboptimal or ineffective care. Errors in recommendations can occur due to several reasons. First, the evidence supporting a particular recommendation may simply be *lacking*, as there are few areas in medicine that are guided by a large number of high quality studies. Second, the supporting studies leading to the recommendation may be *misleading* because of limitations in study designs that result in biased results, and thus incorrect recommendations. Last, evidence may be *misinterpreted* by the guideline authors, due to differences in opinion and values, or due to conflicts of interest.

Evidence-based guidelines attempt to address many of these limitations by critical appraisal and grading of the recommendations. In the ACCP Grading system, data that are lacking is reflected in uncertain recommendations (Grade 2) and study designs that are more prone to bias receive Grade B or C recommendations. The process acknowledges that factors other than literature evidence will influence recommendations—these may be the author's own values and preferences, or conflicts of interest. Consequently, the participants in most guidelines are asked to disclose real or potential financial and intellectual conflicts of interest including honoraria or research funding obtained in the previous two years, or stocks held from companies that may benefit from the guideline recommendations.<sup>11</sup>

There are limitations that cannot be addressed through evidence-based guidelines. First and foremost, patient preferences and values are difficult to capture in a guideline, and assumptions of the patients' best interests are done on the patients' behalf by the guideline authors. Consequently, guidelines may not actually address patients' needs. This limitation is implicit in most guideline development efforts. Furthermore, the recommendations are usually generalized to what is best for the average patient and are not necessarily tailored to an individual patient's needs. Guidelines also frequently do not specifically address resources and costs in their recommendations, which may influence individual decisions outside of simply considering efficacy and safety. Last, even with use of evidence-based medical approaches, critical appraisal of studies in most guidelines is performed by individual authors, which may result in inconsistencies in study quality evaluation.

## Summary

In summary, evidence-based guidelines have potential benefits and limitations. Guideline development has evolved to include several steps, including addressing a specific clinical question, conducting a systematic search of the literature, critically appraising the quality of the evidence, and summarizing the evidence while accounting for patients' preferences and values. Introduction of grading systems to address the limitations of traditional opinion-based practice guidelines helps to standardize interpretation of the guidelines and minimize potential harms.

## Acknowledgments

Dr. Crowther holds a Career Investigator award from the Heart and Stroke Foundation of Ontario. Dr. Lim holds a CIHR Mentor-Mentee award. Dr. Arnold is a CIHR New Investigator.

## Disclosures

Conflict-of-interest disclosure: D.M.A. receives research funding from Hoffman LaRoche. R.P.R. is a consultant for Sanofi-Aventis. M.A.C. is a consultant for Artisan Pharma, Aton Pharma, and Bayer; receives research funding from Bayer, Boehringer-Ingelheim, Astra-Zeneca, the Heart and Stroke Foundation of Canada and the Canadian Institutes of Health Research; receives direct unrestricted research support from Sanofi-Aventis and Leo Laboratories; and receives honoraria from Pfizer, Leo Laboratories, Organon, and Novo Nordisk. The other authors declare no competing interest.

Off-label drug use: None disclosed.

## Correspondence

Wendy Lim, MD, MSc, St Joseph's Hospital, 50 Charlton Avenue East, Room L208, Hamilton, Ontario L8N 4A6 Canada; Phone 905-521-6024; Fax 905-540-6568.

## References

1. Arnold DM, Stasi R. Does *Helicobacter pylori* therapy result in a platelet count improvement in adults with immune thrombocytopenic purpura regardless of *H pylori* infection status? ASH evidence-based review 2008. *Hematology Am Soc Hematol Educ Program*. 2008;31-32.
2. Bachanova V, Connors J. How is Hodgkin lymphoma in pregnancy best treated? ASH evidence-based review 2008. *Hematology Am Soc Hematol Educ Program*. 2008;33-34.
3. Haspel R, Hillmen P. Which patients with paroxysmal nocturnal hemoglobinuria (PNH) should be treated with eculizumab? ASH evidence-based review 2008. *Hematology Am Soc Hematol Educ Program*. 2008;35.
4. Rovosky R, Crowther M. What is the evidence for the off-label use of recombinant factor VIIa (rFVIIa) in the acute reversal of warfarin? ASH evidence-based review 2008. *Hematology Am Soc Hematol Educ Program*. 2008;36-38.
5. Shustov AR, Savage KJ. Does high-dose therapy and autologous hematopoietic stem cell transplantation have a role in the primary treatment of peripheral T-cell lymphomas? ASH evidence-based review 2008. *Hematology Am Soc Hematol Educ Program*. 2008;39-41.
6. Clinical practice guidelines: directions for a new program. Washington, DC: National Academy Press; 1990.
7. Woolf SH, Grol R, Hutchinson A, Eccles M, Grimshaw J. Clinical guidelines: potential benefits, limitations, and harms of clinical guidelines. *BMJ*. 1999;318:527-530.
8. Guyatt G, Cook D, Haynes B. Evidence based medicine has come a long way. *BMJ*. 2004;329:990-991.
9. Lockwood D, Armstrong M, Grant A. Integrating evidence based medicine into routine clinical practice: seven years' experience at the Hospital for Tropical Diseases, London. *BMJ*. 2004;329:1020-1023.
10. Guyatt, G, Cook D, Jaeschke R, Paulker SG, Schunemann HJ. Grades of recommendation for antithrombotic agents: American College of Chest Physicians Evidence Based Clinical Practice Guidelines. *Chest*. 2008;133:123-131.
11. Proceedings of the Eighth ACCP Conference on Antithrombotic and Thrombolytic Therapy: evidence-based guidelines. *Chest*. 2008;133:67S-968S.
12. Atkins D, Best D, Briss PA, et al. Grading quality of evidence and strength of recommendations. *BMJ*. 2004;328:1490.
13. Harbour R, Miller J. A new system for grading recommendations in evidence based guidelines. *BMJ*. 2001;323:334-336.
14. Grimshaw JM, Russell IT. Effect of clinical guidelines on medical practice: a systematic review of rigorous evaluations. *Lancet*. 1993;342:1317-1322.