# A Guide to Guidelines for Professional Societies and Other Developers of Recommendations

## Introduction to Integrating and Coordinating Efforts in COPD Guideline Development. An Official ATS/ERS Workshop Report

**Holger J. Schünemann, Mark Woodhead, Antonio Anzueto, A. Sonia Buist, William MacNee, Klaus F. Rabe, and John Heffner; on behalf of the ATS/ERS Ad Hoc Committee on Integrating and Coordinating Efforts in COPD Guideline Development**

**Organizations around the world are recognizing that guidelines should be based on the best available evidence, that the development of recommendations needs to be transparent, and that appropriate processes should be followed. In June 2007, we convened an American Thoracic Society (ATS)/European Respiratory Society (ERS)-sponsored workshop with over 60 representatives from 36 international organizations to provide advice to guideline developers about the required steps and processes for guideline development using the management of chronic obstructive pulmonary disease (COPD) as an example. Following the workshop, participants completed a series of 14 review articles that underwent peer review and incorporated key new literature until June 2011 for most articles in this series. The review articles evaluate the guideline cycle including: priority setting, question formulation, managing conflict of interest, defining appropriate outcomes, stakeholder involvement, grading the quality of evidence and strength of recommendations, integration of values and preferences, considering resource use, reporting of guidelines, implementation, and adaptation. In this Introduction we frame the background and methods of these reviews and provide the key conclusions of the workshop. A summary of the workshop's conclusions and recommendations was published in *The Lancet*. Given the enormous resources that are spent on research and the importance of providing the best guidance to healthcare decision makers, attributing appropriate funds to research syntheses and transparent, independent guidance for the development of evidence-based guidelines is justified. Furthermore, given the immense amount of work that is required, individuals and organizations need to collaborate to achieve the best possible and cost-effective coordination of these efforts.**

**Keywords:** guidelines; evidence-based medicine; professional medical organizations; recommendations

### INTRODUCTION

Organizations, including the World Health Organization (WHO), are recognizing that guidelines should be based on the best available evidence, that the development of recommendations needs to be transparent, and that appropriate processes should be followed. We followed the example of a published series of articles that advised WHO on the best approaches to guideline development, and evaluated the required steps to develop guidelines for professional societies and other guideline developers (1). In June 2007, we convened an American Thoracic Society (ATS)/European Respiratory Society (ERS)/National Heart, Lung, and Blood Institute (NHLBI)-sponsored workshop with over 60 representatives from 36 international organizations, with the purpose of determining the required steps and processes for guideline development using the management of chronic obstructive pulmonary disease (COPD) as an example (Table 1). From this workshop a series of 14 review articles were produced to provide advice to guideline developers (2). In this Introduction, we frame the background and methods of these papers.

### METHODS FOR THE WORKSHOP

The program development committee (PDC)—that is, the authors of this article—used the recently published series of articles that advised WHO on best guideline development to develop a list of topics that were of highest interest to professional societies, using COPD as an example (1). The topics that were considered of greatest relevance covered five areas:

1. Establishing and managing guidelines and guideline panels
2. Conflicts of interest and the funding of guidelines
3. Practical guideline development
4. The patient at the center
5. Ensuring appropriate use of guidelines

These areas are covered in the following 14 articles:

1. Identifying target audiences: who are the guidelines for? (3)
2. Priority setting (prioritize which questions/recommendations need to be made and establish a list of the most important health care problems, patient populations, and interventions) (4)
3. Guideline group composition, group processes (how to achieve consensus), and consultation processes (e.g., post–guideline development review by independent organizations or reviewers) (5)
4. Managing conflicts of interest and the funding of guideline development (6)
5. Deciding what type of evidence and outcomes to include in guidelines (e.g., mortality, hospitalizations, pulmonary function, etc.) (7)

This article is the introduction to "Integrating and Coordinating Efforts in Chronic Obstructive Pulmonary Disease (COPD) Guideline Development," an American Thoracic Society (ATS) and European Respiratory Society (ERS) Workshop Report. This official ATS/ERS Workshop Report was adopted by the ATS Board of Directors, August 2012, and by the ERS Executive Committee, February 2012.

Support for this conference was made possible (in part) by 1R13HL 90485-01 from the National Institutes of Health. The views expressed in written conference materials or publications and by speakers and moderators do not necessarily reflect the official policies of the Department of Health and Human Services; nor does mention of trade names, commercial practices, or organizations imply endorsement by the U.S. Government.

**Proc Am Thorac Soc Vol 9, Iss. 5, pp 215–218, Dec 15, 2012**

**DOI: 10.1513/pats.201208-053ST**
**Internet address: www.atsjournals.org**

**TABLE 1. INVITED ORGANIZATIONS AND SELECTED PARTICIPANTS**

Organizations
- American Academy of Family Physicians (AAFP)
- American College of Physicians (ACP)
- American College of Chest Physicians (ACCP)
- Agency for Healthcare Research and Quality (AHRQ)
- American Heart Association (AHA)
- Latin-American Thoracic Association (ALAT)
- Alpha one foundation
- American College of Cardiology (ACC)
- Asian Pacific Respiratory Society (APSR)
- Allergic Rhinitis and its Impact on Asthma (ARIA) Guidelines
- American Thoracic Society (ATS)
- Basel Institute of Clinical Epidemiology (BICE)
- Canadian Agency for Drugs and Technologies in Health (CADTH)
- Chinese Respiratory Society
- Cochrane Collaboration
- COPD Foundation
- European Respiratory Society (ERS)
- EU funding section
- Forum of International Respiratory Societies (FIRS)
- Global Alliance against Respiratory Disease (GARD)
- Global Initiative for Chronic Obstructive Lung Disease (GOLD)
- Grading of Recommendations Assessment, Development and Evaluation Working Group (GRADE)
- Guidelines International Network (G-I-N, www.g-i-n.net)
- Horten Center for Patient Oriented Research, University of Zurich (Zurich, Switzerland)
- Infectious Disease Society of America (IDSA)
- Institute of Medicine (IOM)
- Institute for Quality and Efficiency in Health Care, Germany (IQWiG)
- Italian National Cancer Institute "Regina Elena" (Rome, Italy)
- International Union Against Tuberculosis and Lung Disease (IUATLD)
- The Knowledge and Encounter Research Unit, Mayo Clinic College of Medicine (Rochester, MN)
- CLARITY Research Group, McMaster University (Hamilton, ON, Canada)
- National Institute for Health and Clinical Excellence, UK (NICE)
- Norwegian Knowledge Centre for the Health Services, Oslo, Norway
- Pan-African Thoracic Society (PATS)
- Society of General Internal Medicine (SGIM)
- World Health Organization (WHO)
- World Organization of Family Doctors (WONCA)

Participants, affiliations
- Elie Akl, SGIM
- Phil Alderson, NICE
- Antonio Anzueto, ATS
- David Atkins Agency for Healthcare Quality and Research
- Hilda Bastian, Institute for Quality and Efficiency in Health Care, Germany (IQWiG)
- Michael Baumann, ACCP
- Lisa Bero, University of Southern California
- Peter Black, APSR
- Elizabeth Boyd, University of Southern California
- Jean Bousquet, ARIA
- Jan Brozek, Italian National Cancer Institute, Rome, Italy
- A. Sonia Buist, ATS
- Jako Burger, Dutch Institute for Healthcare Improvement CBO
- Doug Campos-Outcalt, AAFP
- Francoise Cluzeau, NICE
- Deborah Cook, Society for Critical Care Medicine
- Judy Corn, ATS
- Alvaro Cruz, WHO
- J. Randall Curtis, ATS
- Yngve Falck-Ytter, Cleveland Clinic
- Marilyn Field, Institute of Medicine
- Tom File, IDSA
- Leonardo Fabbri, ERS
- Atle Fretheim, Norwegian Knowledge Centre for the Health Services, Oslo, Norway
- John Heffner, ATS
- Suzanne Hill, WHO
- Suzanne Hurd, GINA
- Richard Irwin, ACCP

(*Continued*)

**TABLE 1. (CONTINUED)**

- Marcia Kelson, NICE
- Regina Kunz, BICE
- Roman Jaeschke, SCCM
- Monika Lelgemann
- Claude Lenfant, GOLD
- William Macnee, ERS
- Atul Malhotra, ATS
- Mark Metersky, ACCP
- Victor Montori, The Knowledge and Encounter Research Unit, Mayo Clinic College of Medicine (Rochester, MN)
- Greg Morosco, NIH
- David Gutterman, ACCP
- Gordon Guyatt, McMaster University (Hamilton, Canada)
- Rogelio Perez-Padilla, ALAT (Latinamerican Thoracic Association)
- Amir Qaseem, American College of Physicians
- Molly Osborne, ATS
- Andy Oxman, Cochrane Collaboration
- Milo Puhan, Horten Center, Zurich
- Klaus Rabe, ERS
- Holger Schünemann, ATS, Cochrane Collaboration
- Deborah Shure, Food and Drug Administration (FDA)
- Gerald Turino, ATS
- Giovanni Viegi, ERS
- John W. Walsh, Alpha One Foundation
- Mark Woodhead, ERS
- Kevin Wilson, Division of Pulmonary and Critical Care, Department of Medicine, Boston University School of Medicine, Boston, Massachusetts; Deputy Editor, Pulmonary, Critical Care, and Sleep Medicine, UpToDate, Waltham, Massachusetts
- Timothy Wilt, American College of Physicians
- Thomas Woitalla, Price Waterhouse CoopersChris van Wheel, WONCA
- Barbara Yawn, AAFP

6. Incorporating considerations of cost effectiveness, affordability, and resource implications in guideline development (8)
7. Synthesis, grading, and presentation of evidence in guidelines (9)
8. Integrating values and consumer involvement in guidelines (10)
9. Stakeholder involvement: how to do it right (11)
10. How to integrate multi-morbidities in guidelines (12)
11. Moving from evidence to recommendations: developing recommendations in guidelines (13)
12. Reporting and publishing guidelines (14)
13. Disseminating and implementing guidelines (15)

**TABLE 2. KEY VISIONS RESULTING FROM THE ATS/ERS GUIDELINE DEVELOPMENT WORKSHOP**

1. Globalize the evidence
2. Focus on questions that are important to patients and clinicians and include relevant stakeholders in guideline panels
3. Conduct collaborative evidence reviews relevant to health care questions and recommendations
4. Use a common metric to evaluate the quality of evidence and strength of recommendations
5. Consider co-morbidities in guideline development
6. Identify ways that help guideline consumers (clinicians, patients and others) understand and implement guidelines using the best available tools
7. Deal with conflicts of interest (COI) and guideline sponsoring in a highly transparent way
8. Support development of decision aids to assist the implementation of value and preference sensitive guideline recommendations.
9. Maintain a Collaboration of International Organizations
10. Examine collaborative models for funding guideline development and implementation

Reprinted by permission from Reference 2.

**TABLE 3. ESSENTIAL ELEMENTS OF AN INTERNATIONAL COLLABORATION TO ENHANCE GUIDELINE DEVELOPMENT AND IMPLEMENTATION IN RESPIRATORY DISEASE**

| |
|---|
| International representation of all relevant stakeholders |
| Ability and resources to convene individuals and organizations who have the necessary expertise and resources |
| Openness to evaluate and adopt new ideas, diverse perspectives, and unique contributions |
| Mechanisms for ensuring that new ideas are heard |
| Commitment to designing funding mechanisms that eliminate or reduce potential for bias in guideline development |
| Willingness to build on successful models of international collaborations |
| Adoption of safeguards to ensure the collaboration's intellectual independence |
| Resources for building a comprehensive and standardized registry of evidence and guideline development efforts |
| Expertise in the methodology of guideline development and implementation to serve as a consultative resource for other organizations |

Reprinted by permission from Reference 2.

14. Adaptation, applicability, transferability, evaluation, and updating of guidelines (16)

Workshop participants were invited on the basis of their expertise and organizational perspective. We asked leaders in the field of guideline methodology and development, together with other workshop participants and colleagues for the workshop, to submit background documents focusing on key questions that were suggested and vetted by the PDC. The writing groups supplemented these questions with their own suggestions. Group leaders were provided with templates for writing their documents. They were asked to update the existing reviews conducted for the WHO when available or focus on other existing reviews that provided details about their development (1). The workshop included small group discussions of writing groups, presentations by group leaders, and large group discussions. Draft articles were discussed at the workshop and revised following the workshop with input from PDC members. The articles underwent peer review, and for most documents the literature search was updated up to June 2011 (details are provided in the documents in this series).

The workshop documents are not full systematic reviews, although authors were asked to provide available evidence and be systematic and transparent about the methods. The search strategies and background articles are described in each article. We asked authors to focus on descriptive literature and examples, in particular those relevant for the management of COPD. Each article begins with a brief introduction to the topic and a listing of the key questions (Table 1 in each of the articles).

## SUMMARY

The Workshop Report includes a series of documents that will inform guideline developers, in particular professional societies, from around the world. This Report follows a summary of the key conclusions of the workshop and the vision of the PDC for implementing the recommendations of the workshop that was published recently in *The Lancet* (2). Implementation of this vision that we summarize in Tables 2 and 3 will depend on obtaining nonprofit funding and continued efforts of individuals and organizations. Given the enormous resources that are spent on research and the importance of providing best guidance to healthcare decision makers, attributing appropriate funds to research syntheses and transparent, independent guidance for the development of evidence-based guidelines is justified. Furthermore, given the immense amount of work that is required, individuals and organizations need to collaborate to achieve the best possible and cost-effective coordination of the efforts.

***Author Disclosure***: H.J.S. was a member of the GRADE Working Group, which received honoraria and lecture fees regarding GRADE. M.W. consulted with GlaxoSmithKline and received lecture fees from Pfizer. A.A. consulted with Astra-Zeneca, Bayer, Boehringer Ingelheim, Forest and GlaxoSmithKline, and served on advisory committees of Boehringer Ingelheim, Dey and GlaxoSmithKline. He received lecture fees from Bayer and research grants from Eli Lilly, GlaxoSmithKline and Pneuma. He served on the GOLD Executive and Scientific committees. A.S.B. was on advisory committees of Altana, Boehringer Ingelheim, GlaxoSmithKline, Merck, Novartis, Pfizer, Schering and Sepracor. She reported grants to the BOLD Initiative Operations Center from Astra Zeneca, Boehringer Ingelheim, Chiesi, GlaxoSmithKline, Merck, Pfizer and Schering. W.M. consulted with GlaxoSmithKline and was on advisory committees of Almirall, Bayer, GlaxoSmithKline and Micromet. He received lecture fees from GlaxoSmithKline, Healthcare Education Services, Healthmatter Communications and Zambon, and grants from Ceremedix, Pfizer and Unilever. He received royalties from Atlas Medical Publishing, B.C. Decker, Clinical Publishing, Healthpress, Hodder Headline, Professional Communications and Taylor & Francis, and travel awards from Astra Zeneca, Boehringer Ingelheim, GlaxoSmithKline and Micromet. K.F.R. and J.H. reported no commercial interests relevant to subject matter. All workshop participants disclosed receipt of honoraria and travel reimbursement from the American Thoracic Society.

## References

1. Oxman AD, Fretheim A, Schunemann HJ. Improving the use of research evidence in guideline development: introduction. *Health Res Policy Syst* 2006;4:12.
2. Schunemann HJ, Woodhead M, Anzueto A, Buist S, Macnee W, Rabe KF, Heffner J. A vision statement on guideline development for respiratory disease: the example of COPD. *Lancet* 2009;373: 774–779.
3. Yawn P, Akl E, Qaseem A, Black P, Campos-Outcalt D. Identifying target audiences: who are the guidelines for? Article 1 in integrating and coordinating efforts in COPD guideline development. An official ATS Workshop Report. *Proc Am Thorac Soc* 2012;9:219–224.
4. Atkins D, Perez-Padilla R, MacNee W, Buist S, Cruz AA. Priority setting in guidelines. Article 2 in integrating and coordinating efforts in COPD guideline development. An official ATS Workshop Report. *Proc Am Thorac Soc* 2012;9:225–228.
5. Kunz RFA, Cluzeau F, Wilt T, Qaseem A, Lelgemann M, Kelson M, Guyatt G, Schünemann HJ. Guideline group composition and group processes. Article 3 in integrating and coordinating efforts in COPD guideline development. An official ATS workshop report. *Proc Am Thorac Soc* 2012;9:229–233.
6. Boyd E, Baumann M, Curtis JR, Field M, Jaeschke R, Osborne M, Schünemann H. Guideline funding and conflicts of interest. Article 4 in integrating and coordinating efforts in COPD guideline development. An official ATS Workshop Report. *Proc Am Thorac Soc* 2012; 9:234–242.
7. Wilt TJGG, Kunz R, MacNee W, Puhan MA, Viegi G, Woodhead M, Schünemann HJ. Deciding what kind of evidence to include in guidelines. Article 5 in integrating and coordinating efforts in COPD guideline development. An official ATS Workshop Report. *Proc Am Thorac Soc* 2012;9:243–250.
8. Hill OL Sr, Falck-Ytter Y, Cruz A, Atkins D, Baumann M, Jaeschke R, Woitalla T, Schünemann HJ. Incorporating considerations of cost-effectiveness, affordability, and resource implications in guideline development. Article 6 in integrating and coordinating efforts in COPD guideline development. An official ATS Workshop Report. *Proc Am Thorac Soc* 2012;9:251–255.
9. Guyatt G, Akl E, Oxman A, Wilson K, Puhan MA, Wilt T, Gutterman D, Woodhead M, Antman E, Schünemann HJ. Synthesis, grading and presentation of evidence in guidelines. Article 7 in integrating and coordinating efforts in COPD guideline development. An official ATS Workshop Report. *Proc Am Thorac Soc* 2012;9:265–261.
10. Kelson MAE, Bastian H, Cluzeau F, Curtis R, Guyatt G, Montori V, Oliver S, Schünemann HJ. Integrating values and consumer involvement in guidelines: the patient at the center. Article 8 in integrating and coordinating efforts in COPD guideline development. An official ATS Workshop Report. *Proc Am Thorac Soc* 2012;9:262–268.
11. Cluzeau FWJ, Kelson M, Corn J, Kunz R, Walsh J, Schünemann HJ. Stakeholder involvement: how to do it right. Article 9 in integrating and coordinating efforts in COPD guideline development. An Official ATS Workshop Report. *Proc Am Thorac Soc* 2012;9:269–273.
12. Fabbri LMBC, Boschetto P, Rabe KF, Turino G, Buist S, Yawn B, Schünemann HJ. How to deal with multi-morbidities in guideline

development. Article 10 in integrating and coordinating efforts in COPD guideline development. An official ATS Workshop Report. *Proc Am Thorac Soc* 2012;9:274–281.

13. Schunemann H, Oxman A, Akl EA, Brozek J, Montori V, Heffner J, Hill S, Woodhead M, Campos-Outcalt D, Alderson P, *et al*. Moving from evidence to developing recommendation in guidelines. Article 11 in integrating and coordinating efforts in COPD guideline development. An official ATS Workshop Report. *Proc Am Thorac Soc* 2012;9: 282–292.
14. Wilson K, Irwin RS, File TM, Schunemann HJ, Guyatt G, Rabe KF. Reporting and publishing guidelines. Article 12 in integrating and coordinating efforts in COPD guideline development. An official ATS Workshop Report. *Proc Am Thorac Soc* 2012;9:293–297.
15. Grimshaw G, Schünemann HJ, Burgers J, Cruz AA, Heffner J, Metersky M, Cook D. Disseminating and implementing guidelines. Article 13 in integrating and coordinating efforts in COPD guideline development. An official ATS Workshop Report. *Proc Am Thorac Soc* 2012;9:298–303.
16. Burgers J, Anzueto A, Black PN, Cruz AA, Fervers B, Graham ID, Hurd S, Metersky M, Woodhead M, Yawn BP. Adaptation, evaluation, and updating of guidelines. Article 14 in integrating and coordinating efforts in COPD guideline development. An official ATS Workshop Report. *Proc Am Thorac Soc* 2012;9:304–340.

# Identifying Target Audiences: Who Are the Guidelines For?

## Article 1 in Integrating and Coordinating Efforts in COPD Guideline Development. An Official ATS/ERS Workshop Report

**Barbara P. Yawn, Elie A. Akl, Amir Qaseem, Peter Black, and Doug Campos-Outcalt; on behalf of the ATS/ERS Ad Hoc Committee on Integrating and Coordinating Efforts in COPD Guideline Development**

***Background*: Professional societies, like many other organizations around the world, have recognized the need to use rigorous processes to ensure that health care recommendations are informed by the best available research evidence. Different clinical practice guidelines addressing the management of the same disease may vary widely in the evidence used and the format of the recommendations, with the result that not all are appropriate for all audiences. This is the first of a series of 14 articles that clinicians, methodologists, and researchers from around the world prepared to advise those developing guidelines in respiratory and other diseases about the potential impact of identifying the target audiences for their clinical practice guidelines.**

***Methods*: In this review we address the following questions. (*1*) Which audiences are interested in a chronic obstructive pulmonary disease (COP)D guideline? (*2*) How many audiences can be addressed in a single COPD guideline? (*3*) What is the purpose of the guidelines? (*4*) Who should be included on the guideline panel? We collected information by searching PubMed and reviewing information from groups that are currently making and using respiratory disease guidelines, as well as from workshop discussions. Our conclusions are based on available evidence, consideration of what guideline developers are doing, and the opinions of those who attended the workshop.**

***Results and Conclusions*: Clinicians desire COPD and other guidelines that are concise, use evidence from practices similar to theirs, and whose authors have expertise in providing care in similar settings and with similar patients. In the case of COPD, barriers to generalists' use of guidelines include lack of awareness of the guidelines, failure to embrace the diagnostic methods as capable of providing definitive confirmation of COPD, and, most importantly, failure of previous guidelines to address the treatment of COPD in the context of the broad range of multiple morbidities that affect most people with COPD. COPD specialists may require guidelines with more details regarding complex COPD management. The purpose of the guidelines may determine the appropriate audience. Guidelines developed to improve care by enhancing education may have a very different audience than guidelines designed to improve care by limiting the scope of practice, punishing noncompliance, or saving money. The purpose will drive dissemination and implementation strategies, but should not influence the methods used to develop a guideline. Clinicians desire guidelines, but data suggest that the current development systems, content, format, and dissemination strategies may need to be altered to fit these audiences. After the purpose and audience are determined, the guideline committee must decide how to fairly address these audiences, which will usually require seeking their input.**

## INTRODUCTION

Different types of organizations and groups develop and use clinical practice guidelines to outline or define the standards of clinical practice for specific diseases (1). Insurers, whether private or public, also develop and use guidelines to summarize or affirm what health care services they do and do not consider reimbursable for specific groups of people. Clinics, hospitals, and emergency care departments have used guidelines for specific health problems to develop clinical pathways or critical care pathways (2, 3).

The audience of a clinical practice guideline varies and depends in part on the topic of the guideline, the purpose of the guideline, and how the guideline defines health and disease (4). Therefore, the target audience should have an influence on the breadth and depth of the guideline content.

Defining the target audience is a critical first step in the process of guideline development. Without knowledge of by whom and how the guidelines are to be used, it is unlikely that the final product will be of value to the intended audience for the projected purpose (4). In addition to the primary audience for a guideline, it is also important to determine which secondary audience may choose to use the guidelines, how they may need to alter or expand the format and content of the guidelines, and how they will interpret the guideline recommendations.

The key decisions regarding the appropriate audience(s) to develop guidelines have received very little attention in the medical literature. In June 2007, the American Thoracic Society (ATS) and the European Respiratory Society (ERS) convened an international workshop of clinicians, methodologists, and researchers from around the world to coordinate efforts in guideline development for patients with chronic obstructive pulmonary disease (COPD) (5). Participants completed the work during the subsequent 4 years to develop a series of recommendations.

This is the first of a series of 14 articles prepared to advise guideline developers about the full spectrum of guideline development. This article begins the series by discussing identification of the target audiences that use clinical practice guidelines.

## METHODS

The authors of this article developed and discussed the key questions (Table 1) by searching PubMed, as well as reviewed information from groups that are currently making and using respiratory disease guidelines. We updated the literature search through June 2011 to highlight work presented since the conference (1, 5–9). In addition, we reviewed the prefaces and

This article is a section of "Integrating and Coordinating Efforts in Chronic Obstructive Pulmonary Disease (COPD) Guideline Development," an American Thoracic Society (ATS) and European Respiratory Society (ERS) Workshop Report. This official ATS/ERS Workshop Report was adopted by the ATS Board of Directors, August 2012, and by the ERS Executive Committee, February 2012.

Funding for this conference was made possible (in part) by 1R13HL 90485–01 from the National Institutes of Health. The views expressed in written conference materials or publications and by speakers and moderators do not necessarily reflect the official policies of the Department of Health and Human Services; nor does mention of trade names, commercial practices, or organizations imply endorsement by the U.S. Government.

**Proc Am Thorac Soc Vol 9, Iss. 5, pp 219–224, Dec 15, 2012**

**DOI: 10.1513/pats.201208-054ST**
**Internet address: www.atsjournals.org**

**TABLE 1. QUESTIONS ADDRESSED REGARDING IDENTIFYING TARGET AUDIENCES IN GUIDELINE DEVELOPMENT**

1. Which audiences are interested in having COPD guidelines? Who currently uses COPD guidelines, which COPD guidelines do they use, and how do they use them?
2. How many audiences can be addressed with a single COPD guideline? How many sets do we need?
3. What is the purpose of the guidelines?
4. Who should be included on the guideline panel? How is this decided and by whom?

*Definition of abbreviation*: COPD = chronic obstructive pulmonary disease.

credentials of the members of guideline panels to attempt to discern how decisions were made in selecting the audience, scope, and panel membership for the existing COPD guidelines. We did not conduct a systematic review ourselves; rather, we used our own judgments to filter what evidence was reviewed for the existing COPD guidelines. The collected information and key questions were used to guide a review and assessment of the topic at the international workshop. Our conclusions are based on the available evidence, consideration of what guideline developers are doing, and workshop discussions.

## RESULTS

The indexed medical literature contains hundreds of references regarding COPD guidelines. Many describe the "use" of those guidelines in different settings. A few discussed the audience intended for those guidelines. Fewer yet studied or discussed the purpose of the guidelines beyond broad generalizations, such as "improving COPD management." In this section, we review some of the more recent publications that address our key questions.

### 1. Which Audiences Are Interested in Having COPD Guidelines? Who Currently Uses COPD Guidelines? Which COPD Guidelines Do They Use, and How Do They Use Them?

Most high-income countries have developed national clinical practice guidelines for patients with COPD, and researchers in several countries have attempted to assess the use and the impact of those guidelines on clinical care (10–21). The methods used to assess guidelines' use and impact have been limited primarily to self-report and vignette studies. No study has directly assessed the actual use of the specific sets of COPD guidelines. In Germany (14) and Belgium (11), primary care physicians and pulmonologists reported that they liked and used COPD guidelines. However, actual care was often not compatible with the guidelines they reported to use in terms of diagnosis (e.g., care showed lack of spirometry use) and certain aspects of treatment (i.e., frequent use of steroids and the limited implementation of pulmonary rehabilitation). While half of Greek primary care physicians reported using the Global initiative for chronic Obstructive Lung Disease (GOLD) guidelines, one third reported simply following the recommendations of their consulting pulmonologist (22). Tsagaraki and coworkers reported that, in Greece, adherence to pharmacotherapy recommended in international and national guidelines increased from 1997 to 2003, but only for lung specialists and not for primary care physicians (18). Polish specialists' self-reported practice was characterized by an underuse of spirometry and an overuse of oral steroids (21). In all of the studies, the conclusions were similar: better dissemination of COPD guidelines is required and will improve disease management, but better implementation of COPD guidelines is also required.

The value of dissemination of guidelines is highlighted by Japanese researchers who, in 2002, (*1*) reported that Japanese physicians who were aware of the Japanese Respiratory Society's COPD guidelines used the guidelines well, and (*2*) concluded that the COPD guideline "should be more widely disseminated" (16). Yet, 3 years later, a clinical vignette study among Japanese physicians reported that care for patients with COPD diverged widely from published practice guidelines (13). Swiss researchers used a prospective observational study to assess actual diagnostic and therapeutic practices and found them poorly concordant with guidelines. Their conclusions were the same, however, stating that "efforts to improve adherence to the Swiss guidelines for the management of COPD should be intensified" (12, 23). Simply increasing dissemination and "intensifying attempts to increase adherence" seem inadequate (17).

To date, little research has addressed the reasons why primary care physicians are unable to integrate existing guidelines into clinical knowledge and practice. Both generalists and specialists appear to have limited understanding of the need for an objective measure (spirometry) to diagnose COPD (24). Do these physicians simply not understand the guidelines or do they not agree with the guidelines? Do the physicians question the validity of the data that was the basis of the guidelines? Have the guidelines been developed without regard for the audience and without attention to office-based implementation strategies and tools?

Twenty-two percent and 14% of U.S. physicians attending two different COPD symposia reported using COPD and asthma guidelines in the management of their patients; the remaining 64% reported either not using or not knowing about any COPD management guidelines (25). However, 76% of the respondents reported that they were interested in using COPD guidelines, and over 90% desired further education when those guidelines become available (25). Nursing professionals also desire guidelines, but express concerns about existing guidelines not including extensive information on issues that nurses believe are the major foci of their care, including health promotion and patient (and family) education about early recognition of symptoms (26).

At an international level, the World Health Organization (WHO) felt that COPD guidelines for primary care physicians were necessary. The WHO primary care guideline panel included mainly primary care physicians and nurses with a few consulting specialists. These guidelines used a very different approach than other specialty-driven COPD guidelines. They began by addressing the respiratory symptoms for which the patient comes to the office, and then moved through a differential diagnosis with specific questionnaires for assessment. The diagnostic assessments were geared to the different levels of resources available to physicians around the world, ranging from physical examination only in lesser-developed health care facilities to pulmonary function tests and imaging of all types in the richest health care environments (27). The impact of this very different format of guidelines has not been studied. However, the desire of primary care physicians and the WHO to develop these guidelines highlights the interest of primary care physicians and nurses in having input into the content, format, and purpose of guidelines (25, 28, 29).

### 2. How Many Audiences Can Be Addressed in a Single COPD Guideline?

***Should we strive for one set of guidelines or at least one set for each major group of health care professionals working with people with COPD?*** A number of national and international COPD guidelines are targeted to multiple users. The target audience of the COPD guidelines from the American College of Physicians is "all physicians" (30). The National Institute for Clinical Excellence (NICE) also states that its guidelines are "for all physicians" (30); this evidence-based guideline makes over 200 recommendations in seven key areas (31). It is unrealistic to assume that a primary care physician can integrate 200 COPD recommendations into practice when COPD represents only one of the many chronic diseases that he or she must treat daily. Instead, guidelines should increasingly be understood as a resource to answer specific questions that arise in clinical practice, rather than a mandate to implement every recommendation in a guideline. This demands that the guidelines be retrievable when those questions are asked. The Japanese Respiratory Society guidelines published in 1999 recommended thin-slice computed tomography for the diagnosis and classification of COPD severity, which clearly identified the guidelines as specialty focused despite the purpose stated in the introduction as being for all physicians (32). Some guidelines fail to identify any specific audience, suggesting that the process of developing the guidelines may not have even considered the target audience (33).

The multiplicity of guidelines across national and international borders may not be explained just by the professional designation of the intended audiences. Language and culture differences among the targeted groups can also explain the desire for multiple guidelines. However, it is not clear whether overcoming the cultural and language barriers would allow a single set of guidelines to be developed for the needs of primary care and specialty care physicians (26), as well as the nonphysician practitioners (2, 14, 26, 32–35).

A first step may be to globally address reasons for regional differences, including the varying degrees of comfort with the diagnosis of COPD and how the diagnosis is made and presented to the patient as well as different national attitudes regarding the value of various classes of COPD drugs (5, 36).

Studies by Glaab and colleagues (14), Tsoumakidou and coworkers (35), and Ferguson (37) highlight the potentially adverse impact of multiple conflicting guidelines. Current authoritative spirometry guidelines from different organizations and countries use conflicting percentages of $FEV_1$/FVC ratios ($FEV_1$/FVC%) to define airway obstruction. For example, if the National Health and Nutrition Examination Survey (NHANES) II percentiles are used, nearly one half of young adults with $FEV_1$/FVC% below the NHANES-III fifth percentile of normal would be misidentified as normal (false negative) because their $FEV_1$/FVC% would be greater than 70% (abnormal misidentified as normal). Conversely, one fifth of older adults with an observed $FEV_1$/FVC% above the NHANES-III fifth percentile had $FEV_1$/FVC% ratios less than 70% (false positive, normal misidentified as abnormal) (38). Different cutoffs and definitions in different guidelines will lead to diverging approaches to managing the disease, as well as confusion and perhaps a lack of regard for guidelines by the users.

***Should guidelines be tailored to audiences that treat patients in general versus specialty practice?*** The very nature of COPD drives the breadth of the audiences that must be addressed in COPD guidelines. COPD is associated with increasing age, smoking, and inflammation. All three of these factors are associated with multiple other chronic illnesses, as well as the many organ system problems associated directly with COPD. While all audiences for COPD guidelines must be familiar with the multiple morbidities in multiple organ systems associated with COPD, the depth and breadth of discussions of those morbidities may need to vary by audience. For example, all health professionals caring for people with COPD must be aware of common nonpulmonary diseases such as depression experienced by up to 40% of people with COPD. However, only generalists may need specific recommendations on the treatment of depression in COPD since few COPD specialists diagnose or treat depression (39).

### 3. What Is the Purpose of the Guideline?

Determining the purpose of a guideline will likely determine the appropriate audience. Guidelines developed to improve care by enhancing education of diagnostic testing or, conversely, the latest use of lung volume reduction may have very different audiences. Guidelines designed to improve care by limiting scope of practice, punishing noncompliance, or saving money will have audiences determined by the purpose and ability of enforcing such restraints (1).

### 4. Who Should Be Included on the Guideline Panel?

It is important to purposefully select the audience when developing COPD guidelines, and another section in this Workshop Report will deal with the guideline panel composition (40). In addition, the selection of the experts to sit on a guideline panel may also affect the appropriateness of the guidelines to different groups. German specialists developed a COPD guideline that required a primary care evaluation of the patient with suspected COPD, but failed to provide recommendations for a stepwise diagnostic work-up. Primary care experts subsequently used the guidelines and their expertise in office-based practice to develop an algorithm for the stepwise diagnosis of COPD, making the guidelines feasible for primary care implementation (41). A combined group of generalists and specialists might have been able to do this work in a single step, avoiding frustration on both sides. By including the target audience in the guideline panel, the guidelines have a better chance of being implemented in the practices they are targeting (42, 43).

But the target audience may require more than just being included in the guideline panel. Italian General Practitioners (GPs) found that guidelines developed by the target audience (GPs) were well received, liked, and used; however, they also found that the guidelines did not substantially alter the clinical progression of patients with COPD, even though some facets of management improved (44). At the end of the trial, the GPs were not sure that implementation of COPD guidelines was an appropriate use of time or resources. This suggests that certain audiences require assurance that guidelines are of value to not only the care process, but also to patient outcomes. A better understanding of the natural history of COPD may have helped the users of the German guidelines (i.e., GPs) identify improved quality of life or functional status as a desirable outcome.

If guidelines are for primary care, then the majority of the panel developing the guidelines should be composed of primary care physicians and nurses (38). The panel should also include some of the specialty physicians and nurses with whom the primary care professionals collaborate and to whom they refer patients requiring additional care for their COPD (e.g., pulmonologists, cardiologists, nutritionists) (24, 27).

If the guidelines are to be used by managed care organizations or policy makers, then the guidelines must address the needs of these groups and use their language (45, 46). A study of the role of quality managers and other system interventions at the U.S. Veterans Administration Hospitals found that organizations can play an important role in providing a supportive climate

to facilitate their clinicians' adherence to guidelines by implementing processes that make it easier to follow the guidelines and culture changes that makes adherence the anticipated result (10, 19). It seems appropriate to involve system change experts familiar with COPD care in COPD guideline panels (47).

Perhaps the guideline groups should also include those who can help measure the impact of guideline development and dissemination. Even when guidelines are implemented, they may not always lead to the expected outcomes (48). For example, one study found that even in hospital patients treated according to guidelines, the length of stay was not statistically different. This demonstrated the importance of prospectively evaluating clinical practice guidelines before recommending them for widespread implementation and determining what outcomes, such as length of stay, are likely to be impacted by the guidelines to be implemented (45).

## DISCUSSION

A number of clinical practice guidelines addressing COPD exist, and most appear to be used by a few groups of physicians some of the time. While a small number of studies actually assess the use of these guidelines, many studies evaluate whether or not care provided is concordant with the published guidelines. Although the concordance of care with the published guidelines may be the explicit result of guidelines use, it may also be the result of other mechanisms, such as community standards of care, pay-for-performance schemes, quality improvement initiatives, clinician systems, or recommendations from consultants. Although guideline developers and promoters desire to have a direct impact on the provision of COPD care (usually given by primary care physicians), the indirect beneficial effect of the other mechanisms should not be ignored. This implies that those who define community standards of care—insurers, quality improvement agents, health administrators, and specialists—become part of the appropriate audience of the guideline and support the recommendations in the guideline.

A small number of studies using physician and nurse self report state that primary care health professionals desire COPD guidelines. But those studies seldom go on to identify what type of guidelines clinicians desire. Extensive work in the area of asthma and preliminary work in COPD have demonstrated that primary care physicians want more than general summaries of evidence. They require specific suggestions and recommendations, as well as tools to integrate those recommendations into their daily practices (49). Of the existing COPD guidelines, only those specifically developed by generalists appear to provide the type of tools that make it possible to rapidly integrate guidelines into practice (1). Future guideline development ventures may need to go beyond simple evidence-based summaries with extensive dissemination plans and move into translation and implementation programs for their COPD guidelines (5).

The breadth of the material included in guidelines must be based on the target audience. COPD guidelines targeting primary care physicians must place COPD in the context of the whole person and address the multiple organ system impact of COPD, the psychosocial impact of major lifestyle changes, and the commonly associated co-morbidities, such as depression and cardiovascular disease. These guidelines need to have breadth with appropriate, but perhaps limited, depth in some specific sections such as surgical interventions or the treatment of very severe COPD.

Many guidelines are now referring to evidence and many use one of several systems of grading evidence, such as GRADE (50). GRADE separately addresses the risk of bias (7) and the applicability of studies (i.e., the appropriateness of the sample or setting) to the question. The latter is called "directness." This approach allows a stepwise evaluation of, for instance, randomized clinical trials that are most often done in controlled specialty settings unlike those present in primary care practice. Results reported for these specialized settings and carefully selected patients may have very poor external validity or generalizability to the "real world" (directness of the best evidence to the question asked). GRADE requires an evaluation of how directly the generated evidence applies to the target population in a separate, but necessary, step.

The selection of an audience should be a very careful and deliberate decision. In most health care systems around the world, primary care physicians continue to provide and manage the care of the majority of people with COPD. The levels of unrecognized disease, the number of people with COPD who continue to smoke, and the continuing rise in early deaths suggest that primary care of COPD could and should be improved. Primary care physicians and teams are therefore appropriate targets for COPD guidelines.

Well-written and appropriately targeted evidence-based COPD guidelines are expensive and time consuming to develop (5–7, 9). Therefore, it is appealing to have a single guideline to serve multiple purposes, including guiding day-to-day diagnosis and management of COPD in office practice, guiding care of COPD exacerbations in office and hospital, offering a framework for measures of quality of care, guiding coverage and payment decisions made by insurers and national health officials, and helping patients determine what they want and need for the management of their disease (5, 8). These multiple purposes could bring several audiences to the guideline development process and must, therefore, be clearly considered when developing guideline content, implementation tools, dissemination strategies, format, and inclusive language. If the guidelines are to be used in multiple countries, recommendations must be appropriate to the many cultural groups and the varying levels of resources in those countries.

The WHO-sponsored International Primary Care Respiratory Group's COPD guidelines targeted primary care practice and included a majority of primary care physicians on its panel. To date, however, most guideline panels have failed to include significant representation from the group they target (5). Few specialists provide care in the primary care setting. In the UK, for example, most specialists are hospital based and have limited experience with outpatient-based care. In the United States, few respiratory specialists care for the multiple conditions often experienced by a single person with COPD, and few specialists in any country deal with all of the time constraints experienced by primary care settings. Therefore, specialists may find it challenging to understand what tools and implementation strategies are required for primary care implementation of recommendations.

In conclusion, clinicians desire guidelines, but data suggest that the current guideline development systems, content, format, and dissemination strategies are not working and may need to be altered to address the desired audiences. Once the audience(s) for a guideline is determined, then the guideline committee must fairly represent these audiences and seek their input. Finally, guideline developers should be aware that guidelines may end up speaking to multiple audiences, including consumers and policy makers, and even when not intended for this purpose, they can end up having a major influence on allocation of resources and interventions for the target disease.

***Author disclosures***: B.P.Y. consulted with Boehringer Ingelheim, GlaxoSmithKline, Novartis and Pfizer, and was on advisory committees of Astra Zeneca, Merck, Pfizer and Schering. She received grants from Astra Zeneca, Boehringer Ingelheim, Merck, Novartis and Pfizer. E.A.A., A.Q., and P.B. reported no commercial interests relevant to subject matter. D.C.-O. received royalties from McGraw-Hill. All workshop participants disclosed receipt of honoraria and travel reimbursement from the American Thoracic Society.

***Acknowledgments***: The authors gratefully acknowledge the contributions of Suzanne Hill to this article.

## References


1. Graham R, Mancher M, Wolman DM, Greenfield S, Steinberg E (editors); Committee on Standards for Developing Trustworthy Clinical Practice Guidelines; Institute of Medicine. Clinical practice guidelines that we can trust (accessed November 2, 2012). Available from: http://www.iom.edu/Reports/2011/Clinical-Practice-Guidelines-We-Can-Trust.aspx
2. New A. Oxygen: kill or cure? Prehospital hyperoxia in the COPD patient. *Emerg Med J* 2006;23:144–146.
3. Tsai TW, Gallagher EJ, Lombardi G, Gennis P, Carter W. Guidelines for the selective ordering of admission chest radiography in adult obstructive airway disease. *Ann Emerg Med* 1993;22:1854–1858.
4. Lacasse Y, Ferreira I, Brooks D, Newman T, Goldstein RS. Critical appraisal of clinical practice guidelines targeting chronic obstructive pulmonary disease. *Arch Intern Med* 2001;161:69–74.
5. Schunemann HJ, Woodhead M, Anzueto A, Buist S, Macnee W, Rabe KF, Heffner J. A vision statement on guideline development for respiratory disease: the example of COPD. *Lancet* 2009;373:774–779.
6. Mlika-Cabanne N, Harbour R, de Beer H, Laurence M, Cook R, Twaddle S. Sharing hard labour: developing a standard template for data summaries in guideline development. *BMJ Qual Saf* 2011;20:141–145.
7. Steinert T, Richter D, Bergk J. Sources of bias in guideline development. *Psychiatr Serv* 2010;61:946–947.
8. Sanger S, Englert G, Brunsmann F, Quadder B, Villarroell D, Ollenschlager G. Patientenbeteiligung an der Leitlinienentwicklung–sind die Patientenorganisationen fur diese Aufgabe gerustet? [Patient involvement in clinical guideline development: are patient organisations prepared for this task?] *Z Evid Fortbild Qual Gesundhwes* 2009;103:13–16.
9. Cook GE. Guideline development: missing published and unpublished data in guidelines. *BMJ* 2010;340:c1337.
10. Ringbaek TJ, Lange P, Viskum K. Geographic variation in long-term oxygen therapy in Denmark: factors related to adherence to guidelines for long-term oxygen therapy. *Chest* 2001;119:1711–1716.
11. Decramer M, Bartsch P, Pauwels R, Yernault JC. Management of copd according to guidelines: a national survey among Belgian physicians. *Monaldi Arch Chest Dis* 2003;59:62–80.
12. Fritsch K, Jacot ML, Klarer A, Wick F, Bruggmann P, Krause M, Thurnheer R. Adherence to the swiss guidelines for management of COPD: experience of a Swiss teaching hospital. *Swiss Med Wkly* 2005;135:116–121.
13. Fukuhara S, Nishimura M, Nordyke RJ, Zaher CA, Peabody JW. Patterns of care for COPD by Japanese physicians. *Respirology* 2005;10:341–348.
14. Glaab T, Banik N, Rutschmann OT, Wencker M. National survey of guideline-compliant COPD management among pneumologists and primary care physicians. *COPD* 2006;3:141–148.
15. Glaab T, Banik N, Trautmann M, Hellmann A, Wencker M. Guideline-conformity of outpatient COPD management by pneumologists [in German]. *Pneumologie* 2006;60:395–400.
16. Ishii T, Teramoto S, Miyashita A, Suzuki M, Ishigatsubo Y, Kuwahira I, Ueki J, Ouchi Y, Fukuchi Y, Matsuse T. Chronic obstructive pulmonary disease (COPD) in the elderly: analysis from questionnaire about attitudes to the COPD guideline of the Japanese respiratory society and actual therapy for COPD used by physicians [in Japanese]. *Nihon Kokyuki Gakkai Zasshi* 2002;40:113–122.
17. Roede BM, Bindels PJE, Brouwer HJ, Bresser P, de Borgie C, Prins JM. Antibiotics and steroids for exacerbations of COPD in primary care: compliance with Dutch guidelines. *Br J Gen Pract* 2006;56:662–665.
18. Tsagaraki V, Markantonis SL, Amfilochiou A. Pharmacotherapeutic management of COPD patients in Greece: adherence to international guidelines. *J Clin Pharm Ther* 2006;31:369–374.
19. Ward MM, Yankey JW, Vaughn TE, BootsMiller BJ, Flach SD, Watrin S, Doebbeling BN. Provider adherence to COPD guidelines: relationship to organizational factors. *J Eval Clin Pract* 2005;11:379–387.
20. Warren JR, Noone JT, Smith BJ, Ruffin R, Frith P, van der Zwaag BJ, Beliakov GV, Frankel HK, McElroy HJ. Automated attention flags in chronic disease care planning. *Med J Aust* 2001;175:308–312.
21. Zielinski J, Kuca P, Ptasznik P, Plywaczewski R. Management of COPD: pulmonologists' adherence to Polish guidelines [in Polish]. *Pneumonol Alergol Pol* 2005;73:135–141.
22. Trakada G, Spiropoulos K. Chronic obstructive pulmonary disease management among primary healthcare physicians. *Monaldi Arch Chest Dis* 2000;55:201–204.
23. Rutschmann OT, Janssens JP, Vermeulen B, Sarasin FP. Knowledge of guidelines for the management of COPD: a survey of primary care physicians. *Respir Med* 2004;98:932–937.
24. Pearson MG. How can the implementation of guidelines be improved? *Chest* 2000;117:38–41.
25. Yawn BP, Wollan PC. Knowledge and attitudes of family physicians coming to COPD continuing medical education. *Int J Chron Obstruct Pulmon Dis* 2005;3:311–317.
26. Gronkiewicz C, Borkgren-Okonek M. Acute exacerbation of COPD: Nursing application of evidence-based guidelines. *Crit Care Nurs Q* 2004;27:336–352.
27. Bellamy D, Bouchard J, Henrichsen S, Johansson G, Langhammer A, Reid J, van Weel C, Buist S. International primary care respiratory group (IPCRG) guidelines: management of chronic obstructive pulmonary disease (COPD). *Prim Care Respir J* 2006;15:48–57. (see comment).
28. Levy ML, Fletcher M, Price DB, Hausen T, Halbert RJ, Yawn BP. International primary care respiratory group (IPCRG) guidelines: diagnosis of respiratory diseases in primary care. *Prim Care Respir J* 2006;15:20–34. (see comment).
29. van Schayck CP. The revised Dutch college of general practitioners' standard on COPD and the first international who standard: differences and similarities [in Dutch]. *Ned Tijdschr Geneeskd* 2002;146:353–356.
30. Qaseem A, Snow V, Shekelle P, Sherif K, Wilt TJ, Weinberger S, Owens DK. Diagnosis and management of stable chronic obstructive pulmonary disease: a clinical practice guideline from the American College of Physicians. *Ann Intern Med* 2007;147:633–638.
31. Scullion JE. Nice guidelines: the management, treatment and care of COPD. *Br J Nurs* 2004;13:1100–1103.
32. Kitamura S. COPD guideline of Japanese respiratory society [in Japanese]. *Nippon Rinsho* 2003;61:2077–2081.
33. Ollenschlager G, Kopp I, Lelgemann M. The German program for disease management guidelines: COPD guideline 2006. Short review [in German]. *Med Klin* 2007;102:50–55.
34. Pearce L. Putting policy into practice: the example of COPD. *Br J Community Nurs* 2006;11:94–96.
35. Tsoumakidou M, Tzanakis N, Voulgaraki O, Mitrouska I, Chrysofakis G, Samiou M, Siafakas NM. Is there any correlation between the ATS, BTS, ERS and GOLD COPD's severity scales and the frequency of hospital admissions? *Respir Med* 2004;98:178–183.
36. Rudolf M. The reality of drug use in COPD: the European perspective. *Chest* 2000;117:29–32.
37. Ferguson GT. Recommendations for the management of copd. *Chest* 2000;117:23–28.
38. Hansen JE, Sun X-G, Wasserman K. Spirometric criteria for airway obstruction: Use percentage of fev1/fvc ratio below the fifth percentile, not $< 70\%$. *Chest* 2007;131:349–355. (see comment).
39. Norwood R, Balkissoon R. Current perspectives on management of co-morbid depression in copd. *COPD* 2005;2:185–193.
40. Cluzeau FWJ, Kelson M, Corn J, Kunz R, Walsh J, Schünemann HJ. Stakeholder involvement: how to do it right. Article 9 in integrating and coordinating efforts in COPD guideline development. An official ATS workshop report. *Proc Am Thorac Soc* 2012;9:269–273.
41. Schneider A, Borst MM, Gerlach FM, Szecsenyi J. Suspicion of COPD: development of an algorithm for stepwise diagnosis in primary care [in German]. *Z Arztl Fortbild Qualitatssich* 2005;99:217–225.
42. Boulet L-P, Becker A, Bowie D, Hernandez P, McIvor A, Rouleau M, Bourbeau J, Graham ID, Logan J, Legare F, *et al*. Implementing practice guidelines: a workshop on guidelines dissemination and implementation with a focus on asthma and COPD. *Can Respir J* 2006;13:5–47.
43. Meulepas MA, Jacobs JE, Lucas AEM, Smeenk FWJM, Smeele I, Bottema BJAM, Grol RPTM. The feasibility of a primary care model for the management of COPD. *Prim Care Respir J* 2006;15:337–341.
44. Tinelli C, Rezzani C, Biino G, Marinoni A, Rossi A, Volpini E, Pistorio A, Termini R, Grassi M. Evaluation of the efficacy of the Italian guidelines on COPD: a cluster randomized trial. *Monaldi Arch Chest Dis* 2003;59:199–206.
45. Kong GK, Belman MJ, Weingarten S. Reducing length of stay for patients hospitalized with exacerbation of COPD by using a practice guideline. *Chest* 1997;111:89–94.

46. Postma DS. Guidelines and shared care for asthma and COPD. *Neth J Med* 2003;61:63–64.
47. Cannon HE. Assessing the need for a clinical COPD care program in a-managed care organization. *J Manag Care Pharm* 2004;10:S17–S21.
48. Ringbaek TJ. Continuous oxygen therapy for hypoxic pulmonary disease: guidelines, compliance and effects. *Treat Respir Med* 2005;4: 397–408.
49. Yawn B, Bertram S, Wollan P. Introduction of asthma apgar tools improve asthma management in primary care practices. *J Asthma Allergy* 2008;1:1–10.
50. Guyatt GH, Oxman AD, Vist GE, Kunz R, Falck-Ytter Y, Alonso-Coello P, Schunemann HJ, for the GRADE Working Group. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ* 2008;336:924–926.

# Priority Setting in Guideline Development

## Article 2 in Integrating and Coordinating Efforts in COPD Guideline Development. An Official ATS/ERS Workshop Report

**David Atkins, Rogelio Perez-Padilla, William MacNee, A. Sonia Buist, and Alvaro A. Cruz; on behalf of the ATS/ERS Ad Hoc Committee on Integrating and Coordinating Efforts in COPD Guideline Development**

***Introduction*: Professional societies, like many other organizations around the world, have recognized the need to use more rigorous processes to ensure that health care recommendations are informed by the best available research evidence. Priority setting is an essential component of developing clinical practice guidelines informed by the best available research evidence. It ensures that resources and attention are devoted to those areas in which clinical recommendations will provide the greatest benefit to patients, clinicians, and policy makers. This is the second of a series of 14 articles that methodologists and researchers from around the world prepared to advise guideline developers in respiratory and other diseases. This review focuses on priority setting, addressing five key questions.**

***Methods*: In this review, we addressed the following questions. (*1*) At which steps of guideline development should priorities be considered? (*2*) How do we create an initial list of potential topics within the guideline? (*3*) What criteria should be used to establish priorities? (*4*) What parties should be involved and what processes should be used to set priorities? (*5*)What are the potential challenges of setting priorities? We updated an existing review on priority setting, and searched PubMed and other databases of methodological studies for existing systematic reviews and relevant methodological research. We did not conduct systematic reviews ourselves. Our conclusions are based on available evidence, our own experience working with guideline developers, and workshop discussions.**

***Results and Discussion*: Existing literature on priority setting largely applies to identifying priorities for which guidelines to develop rather than setting priorities for recommendations within a guideline. Nonetheless, there is substantial consensus about the general factors that should be considered in setting priorities. These include the burdens and costs of illness, potential impact of a recommendation, identified deficits or weak points in practice, variation or uncertainty in practice, and availability of evidence. The input of a variety of stakeholders is useful in setting priorities, although informal consultation is used more often than formal methods. Processes for setting priorities remains poorly described in most guidelines.**

## INTRODUCTION

Professional societies, like many other organizations around the world, have recognized the need to use more rigorous processes to ensure that health care recommendations are informed by the best available research evidence. Priority setting is an essential component of developing clinical practice guidelines. It ensures that resources and attention are devoted to those areas in which clinical recommendations will provide the greatest benefit to patients, clinicians, and policy makers. Priority setting occurs on three levels. Most attention has been paid to which guidelines should be developed by a sponsoring organization (for example, the World Health Organization or a clinical specialty organization such as the American Thoracic Society [ATS] or European Respiratory Society [ERS]). Similarly, completed guidelines often assign priority to a subset of recommendations of greatest importance for implementation; these may form the basis for quality measures or audit criteria to track performance. Equally important, but less discussed, is the need to set priorities among the many issues that may be addressed within a broader guideline on a specific condition, such as the diagnosis and treatment of chronic obstructive pulmonary disease (COPD). Without some attempt to set relative priorities, sponsors and developers cannot appropriately manage the resources required for the retrieval of evidence, assessment of evidence, or development of recommendations. Initial priority setting can help ensure that the guideline effort devotes sufficient attention to those recommendations that offer the greatest potential to improve health care and health outcomes.

In June 2007 the ATS and the ERS convened an international workshop of methodologists and researchers from around the world for coordinating efforts in guideline development for COPD and other respiratory diseases (1). This is the second of a series of 14 articles prepared to advise guideline developers in respiratory and other diseases on approaches for guideline development. In this article we focus on priority setting.

## METHODS

The key questions for this review were vetted among the authors of this paper (Table 1). We updated a review of the literature on priority setting (2), but we did not conduct a full or formal systematic review. We modified and expanded the questions addressed by that review, to address the five key questions listed in Table 1. We built on the search strategy of a recent review of this issue, which involved a search of PubMed and three databases of methodologic studies relating to guidelines and systematic reviews (the Cochrane Methodology Register, the U.S. National Guideline Clearinghouse, and the Guidelines International Network). We searched PubMed from 1990 to September 2011 using the terms "guidelines" and "priority," and we searched the other databases using the terms "priority," "priority setting," and "topic." We also consulted references from the previous review and our own files. Finally, we reviewed guidelines as of September 2011 on COPD from major international organizations and examined whether they described their process for identifying topics for review and recommendation. Due to the limited literature and dearth of studies describing empirical evidence, the conclusions and recommendations reflect a combination of evidence, the reported practices of organizations involved in developing guidelines, and our own personal experience in priority setting.

---

This article is a section of "Integrating and Coordinating Efforts in Chronic Obstructive Pulmonary Disease (COPD) Guideline Development," an American Thoracic Society (ATS) and European Respiratory Society (ERS) Workshop Report. This official ATS/ERS Workshop Report was adopted by the ATS Board of Directors, August 2012, and by the ERS Executive Committee, February 2012.

Funding for this conference was made possible (in part) by 1R13HL 90485–01 from the National Institutes of Health. The views expressed in written conference materials or publications and by speakers and moderators do not necessarily reflect the official policies of the Department of Health and Human Services; nor does mention of trade names, commercial practices, or organizations imply endorsement by the U.S. Government.

**Proc Am Thorac Soc Vol 9, Iss. 5, pp 225–228, Dec 15, 2012**

**DOI: 10.1513/pats.201208-055ST**
**Internet address: www.atsjournals.org**

**TABLE 1. QUESTIONS FOR PRIORITY SETTING**

1. At which steps of guideline development should priorities be considered?
2. How do we create an initial list of potential questions that need to be prioritized within the guideline development process?
3. What criteria should be used to establish priorities?
4. What parties should be involved and what processes should be used to set priorities?
5. What are the potential challenges in setting priorities?

## RESULTS

There is a limited empirical basis for setting priorities or for addressing the specific questions that we identified. The literature that exists largely applies to identifying priorities among larger and more diverse issues of interest to funding entities or clinical groups. We found no literature on identifying candidates for prioritization and relatively little directly addressing how to prioritize topics within a guideline on a given topic, such as asthma or COPD. Nonetheless, there is substantial consensus about the general factors that can and should be considered in setting priorities for guideline development and a reasonable body of experience with setting priorities within a topic area. The general principles for setting priorities among potential guideline topics also apply to setting priorities among recommendations and topics within a given guideline.

### 1. At Which Steps in Guideline Development Should Priorities Be Considered?

Most of the discussion in the literature has focused on using priorities to decide which guidelines should be developed. We identified, however, a number of additional steps during guideline development at which developers need to weigh priorities (*see* Table 2). The first involves determining which audiences are important for the guideline, as this will dictate the questions and scope of the guideline (3). Once a list of potential questions to be addressed has been prioritized, it is equally important to examine the processes by which one will assemble the evidence to address these questions. There is a growing number of high-quality reviews, especially involving specific medical therapies. The Cochrane Airways Group has over 120 completed reviews of therapies for COPD alone. For many of these topics, especially where strong and consistent evidence exists, an independent review is unlikely to produce new information. Developers should instead focus on areas in which the evidence is most confusing or controversial, in which evidence is rapidly changing, or in which nontrial studies provide essential information (e.g., questions of diagnosis or harms of therapy) and establish ways to incorporate existing reviews to address other questions. Finally, recommendations for practice and research need to be prioritized. This will help direct quality improvement efforts toward the most important changes and will encourage research where it has the greatest potential to improve care.

### 2. How Should We Create a List of Potential Topics within a Guideline?

Most groups leave the task of developing candidate topics within a guideline to the deliberation and consensus of guideline panels. A critical first step in this process is coming to agreement about the target audience for the guideline and the patient populations to which the guideline will apply (3). Guidelines aimed at the pulmonary specialist or in-hospital care will need to consider different topics than those aimed at general practitioners in the outpatient setting. Guidelines addressing common, relatively uncomplicated conditions (e.g., COPD seen in primary care) will have a different set of candidate topics than those that intend to encompass rarer or more complicated variants of disease (e.g., mixed obstructive/restrictive disease due to occupational exposure). Clinical guidelines typically structure the specific content areas to be addressed using the pathophysiology of the disease, different elements of care (e.g., primary and secondary prevention and acute management), and the epidemiology of disease. Most of the available guidelines on COPD are organized around the following broad categories: assessment, diagnosis, monitoring, management of stable disease (including educational, environmental, and pharmacologic and nonpharmacologic interventions), and management of exacerbations (including assessment as well as outpatient and inpatient management).

A variety of methods have been used to identify more specific topics within these areas. Clinicians, experts, and patients can be surveyed for candidate topics. More commonly, guideline panelists and staff use formal or informal processes (including review of other guidelines) to create a list or outline of topics to be covered in a guideline. Increasingly, public entities such as the National Institute for Health and Clinical Excellence in the United Kingdom and the Agency for Healthcare Research and Quality in the United States have developed formal processes to allow stakeholders to comment on the scope and specific questions to be addressed by a systematic review or guideline. Some process to identify "horizon" issues arising from new and emerging technologies and treatments can also be helpful; this can include reviewing abstracts of major research meetings, editorials, and recent drug approvals (4). Generating a candidate list of topics is more straightforward in efforts to coordinate or harmonize existing guidelines, since the original guidelines can be used to generate a list of specific recommendations within each guideline.

**TABLE 2. STEPS FOR SETTING PRIORITIES IN GUIDELINE DEVELOPMENT**

| Priority-Setting Steps in Guidelines | Factors to Consider |
|---|---|
| 1. Identify priorities for guideline development | Consider resources, health burden, stakeholder input, availability of data, variation in care. |
| 2. Indentify target audience and scope of guideline | Consider patterns of care, clinician interest, improvable quality gaps. Adapt scope to address most common problems for target audience. |
| 3. Prioritize questions of potential interest | Consider stakeholder input, epidemiology and costs of disease, practice patterns. Focus on high burden, availability of evidence or ongoing controversy, known practice variation, and potential to change practice. |
| 4. Prioritize effort of synthesizing evidence | Use existing reviews where high-quality reviews are available; focus new effort on those areas with most complicated or controversial evidence. |
| 5. Prioritize recommendations | Focus on recommendations with biggest health impact and best evidence, with emphasis on areas with existing quality gaps. Consider quality measures based on feasibility and reliability of data collection, including adequate sample size. |
| 6. Prioritize recommendations for research | Focus on areas where studies are feasible, would address important knowledge gaps, and have potential to alter practice. |

### 3. What Criteria Should Be Used to Establish Priorities?

There is fairly broad consensus among groups developing guidelines regarding the criteria that can be used to develop priorities. These criteria are generally applicable to setting priorities for recommendations within a specific guideline. The U.S. Institute of Medicine (IOM) Committee on Setting Priorities for Guideline Development outlined six general criteria: prevalence, burden of illness, cost of management, practice variation, potential of a guideline to improve health outcomes, and potential of guideline to reduce costs. Other groups have incorporated similar criteria, but have also included additional criteria: clinical uncertainty or complexity, the availability of adequate evidence/state of knowledge (which should not be influenced by for-profit interest), practitioner interest, and the cost of developing the guideline. In a survey of 55 Canadian organizations involved in guideline development (including researchers, federal and provincial governmental organizations, clinical organizations, industry, and consumers), there was broad agreement that the health burden, economic burden, and state of scientific knowledge were important, while practitioner interest and cost of development were felt to be less important.

Assessing barriers to care may also shed light on where guidelines might help overcome barriers (through educational function or promoting policy changes). A survey of Australian general practitioners identified the education of patients and professional education as the major priorities for optimal asthma care, but identified the time and costs of education and medication as barriers (5).

### 4. What Parties Should Be Involved and What Processes Should Be Used to Set Priorities?

We found a paucity of data and nothing in published COPD guidelines that describe an explicit process for prioritizing specific topics/recommendations *within* a given guideline. The literature on processes for setting priorities comes largely for setting priorities *among* possible guidelines. This literature identifies a variety of stakeholders and end-users that could be involved in the prioritization process. These include clinicians, professional organizations, policymakers, payers (e.g., health plans), government bodies, quality organizations, and patients or patient representatives. Guideline panels, which ultimately set priorities, frequently include representatives of these groups. Although industry is not always included among these stakeholders, government-funded organizations increasingly are allowing input on key questions from a wide array of stakeholders, including industry. Whether or not research conducted by these other stakeholders is included in a guideline should depend on an independent assessment based on transparent inclusion and exclusion criteria.

Recommended processes for setting priorities vary. Batista and Hodge suggest consulting with stakeholders, considering feasibility, and documenting process for setting priorities (6). The IOM recommended expert consultation using explicit questions and use of Delphi-like processes (7). In our personal experience, we have relied on less formal processes. The United States Preventive Services Task Force (USPSTF) has used more formal processes in the past to rank potential prevention topics for developing or updating guidelines; these included surveying experts, assessing criteria such as burden of disease, and limited literature searches to assess current controversy. Because this process proved resource intensive and many criteria (such as variation in practice and potential impact of a guideline) were largely subjective, the USPSTF now sets priorities using similar criteria, but relies on the expert judgments of the panel (8) (Mary Barton, AHRQ, personal communication).

The Evidence-based Practice Centers of the Agency for Healthcare Research and Quality sets priorities within each of its broader systematic reviews with input of a small technical expert panel (9). The panels usually include clinical experts, specialist society representatives, researchers, and often funders or policy makers. Clinical and policy importance of the questions and availability of evidence are usually major considerations.

There is no clear consensus on how to involve patients and the public in setting priorities. An increasing number of guideline panels include patients and some processes allow for public and patient comment at various stages of developing a guideline or review. Patient input is important because patients may place a higher value than clinicians on certain issues addressed within a guideline (e.g., the content of patient education and risk communication).

### 5. What Are the Potential Challenges of Setting Priorities?

The major challenges to setting priorities are the lack of data on many of the common criteria proposed to set priorities, the time involved in formal priority setting, and possible competing perspectives of different stakeholders. Objective data on criteria, such as variation in practice or importance to clinicians, is often lacking. Even when it exists (e.g., through the reporting of performance on quality indicators, such as appropriate prescriptions of controller medications in patients with asthma), it is often unclear whether the variation is due to barriers in knowledge that guidelines can address, or due to structural barriers such as lack of access to care. Equally difficult to assess is the potential for a guideline (or a specific recommendation within a guideline) to improve outcomes or reduce costs. Although it is possible to collect evidence to try to inform some of these issues, it is not clear whether the effort to produce more objective measures of disease burden, gaps in care, and costs lead to different priorities than informed, but subjective, judgment. Because priority setting is more subjective than arriving at recommendations for which there is more robust evidence, panels should be aware of how priorities are influenced by the different perspectives of individual guideline panelists, and their conflicts must be known (10, 11). For example, specialists and researchers may place a high priority on issues that are less relevant to generalists, and vice versa.

## DISCUSSION

Explicit evidence-based clinical practice guidelines were initially developed as tools to improve health outcomes and to reduce healthcare costs. The variety of stakeholders for guidelines has grown, however, as have the intended aims of guidelines. Guidelines now serve as educational resources for clinicians and patients, provide a reference point for reimbursement decisions and quality improvement activities, and serve to draw the attention of the public, policy makers, funding bodies, and researchers to specific issues. It is a challenge to meet all these needs within the resources usually available for developing a guideline. Thus, a clear and coherent process is needed to prioritize which issues will be addressed and how they will be addressed, as well as to identify where to devote resources for a more rigorous and systematic review of the evidence.

Guideline developers may feel compelled to address a wide range of issues out of a desire to be a stand-alone resource or educational tool. It is inefficient and impractical, however, to devote the same level of effort to addressing each of the potential issues in a single guideline. Setting priorities allows developers to focus their greatest attention (e.g., systematic review of the primary literature) on those topics for which the evidence is confusing, practice is variable, recommendations are controversial, or the issues account for a large proportion of the morbidity of the

**TABLE 3. CRITERIA FOR SETTING PRIORITIES**

| Criteria for Setting Priorities | Source of Data |
|---|---|
| Burden of disease (health or economic) | National data; review articles |
| Costs of care | National data; review articles |
| Variability in practice | Quality measurement; surveys; expert opinion |
| Potential impact of guideline or recommendation | Expert opinion and stakeholder input |
| Importance to clinicians | Survey, consultation, and *ad hoc* stakeholder input |
| Importance to patients | Survey, consultation, and *ad hoc* stakeholder input |
| Availability of evidence | Existing reviews; preliminary literature search |
| Uncertainty or controversy | Literature search for editorials, |
| Emerging issues | Meetings, drug/device approvals, policy experts |

condition. Other topics that are lower priority can often be addressed by relying on existing reviews or through informal consensus if there is little evidence.

Guideline developers appear to have gravitated toward generally similar approaches for setting priorities. This includes consulting with stakeholders, considering common criteria (e.g., burden of disease and potential impact), and relying on panels or panel subcommittees to develop priorities. It is not clear that the effort to be involved in more formal processes to set priorities would lead to appreciably different or better outcomes. Nonetheless, the priority setting process should not be neglected within the methods of guideline documents.

## Conclusions

Priority setting should begin with a clear statement of the intended audience, patient population, and outcomes of interest for a guideline. Once these parameters are established, potential issues of interest for a guideline for COPD should be organized into three broad areas (Table 3): (*1*) assessment, diagnosis, and monitoring; (*2*) management of stable disease (including educational, environmental, and pharmacologic and nonpharmacologic interventions); and (*3*) management of exacerbations (including assessment as well as outpatient and in-patient management). Panels should specify factors to be considered in setting priorities for making recommendations, but they should also separately identify those topics for which an independent systematic review of the evidence is justified. The priority-setting process may include some initial guidance about which literature should be reviewed, or those decisions may be left up to the experts conducting the review. The input of policy makers, patients, and researchers, as well as clinical experts, is important in setting priorities. This input can be collected by contacting individual representatives of these different groups, by allowing public comment on priorities, or through formal or informal group processes. We recommend some processes to identify issues that may be priorities due to emerging technology, current policy debates, or evolving practice.

***Author Disclosures***: D.A. and R.P.-P. reported no commercial interests relevant to subject matter. W.M. consulted with GlaxoSmithKline and was on advisory committees of Almirall, Bayer, GlaxoSmithKline and Micromet. He received lecture fees from GlaxoSmithKline, Healthcare Education Services, Healthmatter Communications and Zambon, and grants from Ceremedix, Pfizer and Unilever. He received royalties from Atlas Medical Publishing, B.C. Decker, Clinical Publishing, Healthpress, Hodder Headline, Professional Communications and Taylor & Francis, and travel awards from Astra Zeneca, Boehringer Ingelheim, GlaxoSmithKline and Micromet. A.S.B. was on advisory committees of Altana, Boehringer Ingelheim, GlaxoSmithKline, Merck, Novartis, Pfizer, Schering and Sepracor. She reported grants to the BOLD Initiative Operations Center from Astra Zeneca, Boehringer Ingelheim, Chiesi, GlaxoSmithKline, Merck, Pfizer and Schering. A.A.C. was employed by the World Health Organization. All workshop participants disclosed receipt of honoraria and travel reimbursement from the American Thoracic Society.

## References

1. Schunemann HJ, Woodhead M, Anzueto A, Buist S, Macnee W, Rabe KF, Heffner J. A vision statement on guideline development for respiratory disease: the example of COPD. *Lancet* 2009;373:774–779.
2. Oxman AD, Schunemann HJ, Fretheim A. Improving the use of research evidence in guideline development: 2. Priority setting. *Health Res Policy Syst* 2006;4:14.
3. Yawn P, Akl E, Qaseem A, Black P, Campos-Outcalt D. Identifying target audiences: who are the guidelines for? Article 1 in integrating and coordinating efforts in COPD guideline development. An official ATS Workshop Report. *Proc Am Thorac Soc* 2012;9:219–224.
4. Douw K, Vondeling H. Selection of new health technologies for assessment aimed at informing decision making: a survey among horizon scanning systems. *Int J Technol Assess Health Care* 2006;22: 177–183.
5. Goeman D, Hogan CD, Aroni RA, Abramson MJ, Sawyer SM, Stewart K, Sanci LA, Douglass JA. Barriers to delivering asthma care: a qualitative study of general practitioners. *Med J Aust* 2005; 183:457–460.
6. Battista R, Hodge MJ. Setting priorities and selecting topics for clinical practice guidelines. *CMAJ* 1995;153:1233–1236.
7. Committee on Methods for Setting Priorities for Guidelines Development, Institute of Medicine. Setting priorities for clinical practice guidelines. Field MJ, editor. Washington, D.C.: National Academy Press; 1995.
8. Harris RPHM, Woolf SH, Lohr KN, Mulrow CD, Teutsch SM, Atkins D; Methods Word Group. Third US Preventive Services Task Force. Current methods of the US Preventive Services Task Force: a review of the process. *Am J Prev Med* 2001;20:21–35.
9. Atkins D, Fink K, Slutsky J. Better information for better health care: the evidence-based practice center program and the agency for healthcare research and quality. *Ann Intern Med* 2005;142:1035–1041.
10. Boyd E, Baumann M, Curtis JR, Field M, Jaeschke R, Osborne M, Schünemann H. Guideline funding and conflicts of interest. Article 4 in integrating and coordinating efforts in COPD guideline development. An official ATS Workshop Report. *Proc Am Thorac Soc* 2012; 9:234–242.
11. Schunemann HJ, Osborne M, Moss J, Manthous C, Wagner G, Sicilian L, Ohar J, McDermott S, Lucas L, Jaeschke R. An official American Thoracic Society policy statement: managing conflict of interest in professional societies. *Am J Respir Crit Care Med* 2009;180:564–580.

# Guideline Group Composition and Group Processes

## Article 3 in Integrating and Coordinating Efforts in COPD Guideline Development. An Official ATS/ERS Workshop Report

**Regina Kunz, Atle Fretheim, Françoise Cluzeau, Timothy J. Wilt, Amir Qaseem, Monika Lelgemann, Marcia Kelson, Gordon Guyatt, and Holger Schünemann; on behalf of the ATS/ERS Ad Hoc Committee on Integrating and Coordinating Efforts in COPD Guideline Development**

***Background*: Professional societies, like many other organizations around the world, have recognized the need to use more rigorous processes to ensure that health care recommendations are informed by the best available research evidence. This is the third of a series of 14 articles that were prepared to advise guideline developers in respiratory and other diseases on considerations for group compositions and group processes in guideline development, and how this can be effectively integrated in the context of respiratory disease guidelines on a national and international level.**

***Methods*: We updated a review of the literature addressing group composition and group process, focusing on the following questions:**

1. **How to compose a functioning and representative guideline group**
   - **Who should be included in a guideline panel?**
   - **How to select organizations, groups, and individuals**
   - **What expertise is needed?**
   - **Consultation with nonincluded groups**
2. **How to assure a functioning group process**
   - **How to make the process constructive**
   - **Balancing participation and finding agreement**
   - **Administrative support**
   - **What constitutes sufficient resources?**

**Our conclusions are based on available evidence from published literature, experience from guideline developers, and workshop discussions.**

***Results and Conclusions*: Formal studies addressing optimal processes in developing guidelines are limited, and experience from guideline organizations supplement the formal studies. When resources are available, guideline development groups should aim for multidisciplinary groups, including patients. Prerequisites for a multidisciplinary group include: a strong chair experienced in group facilitation with broad acceptance in the group, training the group in guideline methodology, and professional technical support. Formal consensus developing methods have proved effective in reaching agreement on the final recommendations.**

## INTRODUCTION

Professional societies, like many other organizations around the world, have recognized the need to use more rigorous processes to ensure that health care recommendations are informed by the best available research evidence. A thoughtful composition of a guideline panel and functioning group processes are pivotal for the development of optimal clinical guidelines. In June 2007 the American Thoracic Society (ATS) and the European Respiratory Society (ERS) convened an international workshop of methodologists and researchers from around the world to coordinate efforts in guideline development using chronic obstructive pulmonary disease (COPD) as a model. Participants completed the work during the subsequent 4 years to develop a series of recommendations. This is the third of a series of 14 articles prepared to advise guideline developers in respiratory and other diseases on considerations for group compositions and group processes in guideline development, and how this can be effectively integrated in the context of respiratory disease guidelines on a national and international level.

In this article we addressed the questions listed in Table 1. This section complements two other sections published in this issue: one on stakeholders (1) and one on the involvement of patients and caregivers (2).

## METHODS

The authors of this article developed and discussed the key questions in this article. We updated a review of the literature addressing group composition (3) and group process (4), focusing on the key questions above. We searched PubMed and other databases of methodological studies for existing systematic reviews and relevant methodological research up to June 2011. We did not conduct systematic reviews ourselves. Our conclusions are based on available evidence from published literature, experience from guideline developers, and workshop discussions.

## RESULTS

### How to Compose a Functioning and Representative Guideline Panel

***Should guideline panels include multiple disciplines?*** Developing guidelines based on evidence syntheses often involves disagreement, incorporating viewpoints and preferences of multiple stakeholders, and negotiation. Composition of the guideline panel can impact its recommendations (3). Groups funded by organizations with direct responsibilities to the public (such as the National Institute for Health and Clinical Excellence [NICE]) recruit panels with a broad representation of stakeholders much more often than those convened by professional societies (5–10), whose guidelines are targeted primarily at health professionals. Multidisciplinary groups composed of health care providers across the

---

This article is a section of "Integrating and Coordinating Efforts in Chronic Obstructive Pulmonary Disease (COPD) Guideline Development," an American Thoracic Society (ATS) and European Respiratory Society (ERS) Workshop Report. This official ATS/ERS Workshop Report was adopted by the ATS Board of Directors, August 2012, and by the ERS Executive Committee, February 2012.

Funding for this conference was made possible (in part) by 1R13HL 90485–01 from the National Institutes of Health. The views expressed in written conference materials or publications and by speakers and moderators do not necessarily reflect the official policies of the Department of Health and Human Services; nor does mention of trade names, commercial practices, or organizations imply endorsement by the U.S. Government.

H.J.S. is supported by a European Commission: The human factor, mobility and Marie Curie Actions Scientist Reintegration Grant (IGR 42192) "GRADE."

**Proc Am Thorac Soc Vol 9, Iss. 5, pp 229–233, Dec 15, 2012**

**DOI: 10.1513/pats.201208-056ST**
**Internet address: www.atsjournals.org**

**TABLE 1. QUESTIONS ADDRESSED REGARDING GROUP COMPOSITION AND PROCESS IN GUIDELINE DEVELOPMENT**

| |
|---|
| 1. How to compose a functioning and representative guideline panel |
| Who should be included in a guideline panel? |
| How to select organizations, groups, individuals |
| What expertise is needed? |
| Consultation with nonincluded groups |
| 2. How to assure a functioning group process |
| How to make the process constructive? |
| Balancing participation and finding agreement |
| Administrative support |
| What constitutes sufficient resources? |

clinical care spectrum (primary care and specialists) and from a diverse geographic spread, patients, epidemiologists and health services researchers, health care managers, and so on offer a broader view on health care issues and some protection against dominance by a single group's agenda, and can balance individual biases (11). Depending on the disciplines and the determinants of dissemination, multidisciplinary involvement may generate a sense of ownership that facilitates adoption and implementation of the guideline (12). The section by Cluzeau and colleagues in this issue describes the potential participants of guideline panels including caregivers, patients, employers, manufacturers (pharmaceutical companies), the healthcare industry, and others involved in receiving care, managing care (policy makers, public health services), monitoring care (quality assurance companies), and financing care (governments, health insurers, the public) (1).

The results from studies regarding the impact of multidisciplinary involvement are not entirely consistent. One study reported that dialogue between members of various professional disciplines tends to increase divergence from the evidence they had appraised when formulating their recommendations (13). Not all studies have supported a positive impact of multidisciplinary involvement on dissemination and implementation (14). However, it is clear that the expertise required to develop clinical practice guidelines (e.g., training in critically assessing and summarizing evidence, as well as leading guideline panels) differs from that of clinician experts and clinical researchers.

EXPERTS MAY DOMINATE A GUIDELINE DEVELOPMENT GROUP. The management of complex and chronic diseases often involves several specialties, while the main responsibility (and contact time) remains with the primary care physicians. Studies on group dynamics in guideline development groups (GDGs) consistently report a clear relationship between status, contribution, and impact on decision making (15). Not only have experts and consultants a higher status, they frequently outnumber the primary care physician(s) (16). Adequate representation in the GDG (e.g., more than one primary care physician) and special quotas or exclusion during votes on recommendations can counteract this imbalance.

IS MULTIDISCIPLINARY INVOLVEMENT PRACTICAL? It may be impossible to include representatives of all relevant constituencies (e.g., patients and primary care providers) and still achieve a manageable group size. The inclusion of a chair, professionals, patients, and technical support quickly reach this size, and national guidelines easily attract the interest of 15 or more professional organizations. For example, the German National Guideline on Diabetes (17) involved 28 different societies with a stake in the guidelines. Groups with increasing size are more difficult to manage in terms of debate and decision-making; some evidence suggests that problems arise when there are more than 15 members (18, 19). Increasing size to include all relevant groups may create efficiency problems without improving outcome (20). If GDGs choose a comprehensive multidisciplinary approach, they need to secure additional resources for organizational, administrative, and logistic support. Another approach is to confine the GDG to a core group and offer other stakeholders, including the general public, the opportunity to comment on draft versions of the guideline. For example, NICE uses these consultation periods after making initial drafts of guidelines available to stakeholders.

***How to select organizations, groups, and individuals.*** In general, scientific societies, institutions of the health care system, or professional bodies that prompt the development of a guideline ultimately decide the composition of a GDG. Explicit criteria regarding who should participate (and why) increase the credibility of the selection process. NICE, for example, has decided to openly advertise chair and positions on the GDG and then to select applicants with an explicit job/person profile. The limited number of GDG positions inevitably excludes certain stakeholders if the interest exceeds the number of available positions. The all-inclusive approach of programs such as the National Disease Management/Guideline Program in Germany (21) demands skilled management of the GDG by delegating specific tasks to subgroups, while a steering committee overlooks the whole process and resumes the results.

CONFLICT OF INTEREST. Experts invariably have potential conflicts of interest (COI) aside from financial conflicts, the definition of which is challenging. Initial unrevealed conflicts of interests that ultimately receive public attention can damage the credibility of a guideline (22). Organizations increasingly insist on a declaration of financial conflicts from all panelists that they share with other GDG members. Some organizations, including NICE, insist that conflicts of interest be published with the guideline. Furthermore, some organizations request that members abstain from voting when a recommendation involves conflict of interest. The section by Boyd and coworkers in this series describes guidance for professional societies in dealing with conflict of interest (23).

ROLE OF INDUSTRY. Much dispute concerning the role of stakeholders with commercial interests exists. Should they be explicitly excluded from multidisciplinary groups? NICE, along with most guideline producers, exclude the pharmaceutical industry or manufacturers of devices from the active development of guidelines because it believes their commercial interest would be too significant a conflict of interest to ensure their objective contribution to the guideline development group and it would affect the users' perception of the guideline to such an extent that it would be undesirable. However, most GDG groups invite industry to participate during the public review process on the scope and drafts of the guideline. This review, however, should focus on errors of fact in the guideline such as inadequate dosing and information about study design and conduct, but should not influence the development of recommendations within a guideline. The section by Cluzeau and colleagues in this issue describes this topic in more detail (1).

***What expertise is needed?*** Producing evidence-based practice guidelines requires the participation of individuals with skills beyond being an expert on the topic. A GDG requires expertise from health care professionals, input from patients about their needs and preferences, as well as methodologists and librarians with expertise in gathering, summarizing, and interpreting the evidence. Given the increasing demands on and expertise required to manage complicated guideline development projects, structured education in this methodological area will be required. An international survey of 18 guideline organizations and professional societies reported that most guideline programs offer training in guideline methodology and many plan to increase training in the near future to improve the quality of the guideline (5).

Systematically assessing the literature and producing evidence tables constitutes the bulk of the work for a guideline group. Some organizations regard this work as a core task of the content experts (24), while others commission professional systematic reviewers

**TABLE 2. FORMAL CONSENSUS TECHNIQUES MODIFIED FROM REFERENCE 20**

| Consensus Development Method | Mailed Question naires | Private Decisions Elicited | Formal Feedback of Group Choices | Face-to-Face Contact | Structured Interaction | Aggregation Method |
|---|---|---|---|---|---|---|
| Informal | — | — | — | + | — | Implicit |
| Delphi method | + | + | + | _ | + | Explicit |
| Nominal group Technique | — | + | + | + | + | Explicit |
| Consensus development conference | — | — | — | + | — | Implicit |

to conduct the literature review and synthesis that the GDG uses as the basis for discussion (18, 25). The choice of approach involves the methodological expertise of the panel members, their time commitment, the complexity of the medical problem, and the resources available. Although not every panel member needs to have a profound understanding of guideline methodology, every panel member should have at least some.

***Consultation with nonparticipating groups.*** Ideally, GDGs will establish, from the outset, a system that facilitates input from stakeholders, including the public and industry, accompanied by a policy on how the GDG will deal with the feedback. This is much more realistic for governmental bodies than for professional societies, but it is potentially a goal toward which professional societies can strive. Widespread consultation offers those not included in the GDG the opportunity to comment on the draft, to identify additional evidence, to query judgments, and point to consequences that the panel has not considered. Placing the draft on the internet has become a time-efficient avenue for eliciting feedback (5) but, in addition, the GDG should contact key stakeholder organizations directly. Expert authorities may also provide valuable feedback.

Optimal timing of input remains uncertain. Many organizations ask for input to the final draft of the guideline. Often, however, only major errors or omissions can be corrected at that stage (5). More promising is the current trend to collect public feedback at various milestones in the development process (6, 7). Possible milestones include the decision regarding the scope of the guideline, when the evidence summarization process is complete, and at the time of the first and then the penultimate draft. Since guidelines require revisions, setting up a discussion forum after its publication can collect valuable feedback directly from the guideline users.

### How to Ensure Effective Group Functioning

***How to make the process constructive.*** Functioning groups are prerequisite to successful guideline development. The chair has a crucial role in accomplishing this goal and should, therefore, be experienced in group facilitation, maintaining constructive dynamics, and identifying conflicts early. The chair should also have a repertoire of active conflict management (26). The chair needs acceptance in the group and the authority to enforce any ground rules. At the same time, the chair needs a sufficient understanding of the medical content. In some GDGs, the chair is responsible for leading the group and primarily responsible for drafting the recommendations. In other groups, the chair is primarily a facilitator who takes a neutral position toward all group members, while ensuring balanced participation. Expertise in guideline methodology and continuous interaction with the technical team can support the chair's leadership. Co-chairing of a second individual to complement experience is another alternative to ensure optimal group processes.

***Balancing participation and reaching agreement.*** Panel member input will vary depending on role and status within the group, background, and type of content expertise (such research, practice, politics, and process). Experts tend to be the most active participants in many groups (16) and typically voice their views on research and politics, while primary care physicians, allied health professionals, and the public concentrate their contributions on practice and process (27). Organizations that place a high priority on active participation from patient or public members of the GDG have recognized the need for special support to those members to enable their full participation (16, 28). Systematic exploration of the best approach to ensuring optimal patient participation has just begun.

One of the chair's key functions is to provide panel members with equal opportunities to contribute and to give their arguments appropriate consideration when formulating recommendations. Evidence from a systematic review suggests that formal consensus processes such as Nominal group process or the Delphi method (Table 1) surpass informal processes in achieving agreement (20). The increasing popularity of formal processes in the guideline community suggests their usefulness (5, 18) (Table 2). They prove particularly helpful when only low- or very low–quality evidence is available to answer an important clinical question, as was the case in the NICE guideline on chronic fatigue syndrome where a two-round Delphi survey was established (29). Here, the balanced input and the transparency of the process gave legitimacy to the recommendations.

No single method, however, seems superior to others, and some of them might serve different purposes (e.g., some preferred the Delphi technique in preparing recommendations and background material and the nominal group process for grading recommendations, while others favored the Delphi process for achieving final agreement) (20). Thus, chairs might choose on the basis of their own or the group's familiarity with the technique. While formal consensus methods are a very helpful tool to facilitate discussion, appropriate participation, and agreement, they require substantive administrative support (Table 3).

***Administrative support.*** A GDG's optimal function requires excellent communication and management. Keeping informative minutes to document the process and panel members responsibilities including explicit timelines is helpful. A well-

**TABLE 3. ADVANTAGES AND LIMITATIONS OF FORMAL CONSENSUS PROCESSES**

| Advantages | Limitations |
|---|---|
| Opportunity for active and equal involvement for all participants | Need for a skilled and trained facilitator |
| Potential for definite conclusions as a result | Requires a representative panel to achieve valid recommendations |
| Control of articulate members | Limited by group size (more difficult with large groups) |
| Toning down the power of strong individuals | Can become a mechanical and inflexible tool if poorly moderated |
| Opportunity for panel members to retract from firmly stated opinions without losing face | Can require substantial administrative support |

functioning information management system, such as an intranet system that all members can access, can facilitate structured and timely distribution of documents. Regular contacts with the panel and well-timed reminders keep the guideline commitment a high priority. Many guideline groups suffer as a consequence of insufficient attention to these organizational issues. Thus, proposing a guideline project involves the planning for adequate administrative support.

***What resources are sufficient?*** High-quality guidelines require substantial financial and nonfinancial resources and commitment from panel and staff. Panel members participate for academic credit, recognition as experts, or in support of their medical discipline, but frequently with limited time for guideline work. GDGs that fail to recognize such limitations tend to set overly ambitious goals that result in diminished enthusiasm and unsatisfying compromises on quality.

Starting with a thorough assessment of the proposed objectives, all steps involved, the panel's methodological expertise, and the necessary resources can minimize frustration. This assessment should lead to setting objectives commensurate with the available resources and expertise. Setting realistic objectives may involve concentrating on the most relevant questions (defined by uncertainty about the best management, or practice variation). GDGs may expand resources by sharing work with other groups. If money is available, the GDG may outsource parts of the work to professional reviewers and concentrate the input of the experts to initially formulating questions and ultimately recommendations.

## CONCLUSIONS

Although evidence regarding the best process for the guideline group composition and group processes is limited, some guidance is available. GDGs should strive for wide representation, including patient and public representation, particularly when the issues they address are politically sensitive or when stakeholders are likely to have different perspectives. Therefore, GDGs must balance comprehensive representation against a manageable panel size. A GDG chair with expertise in research methodology, but limited content expertise in the content area, may be primarily responsible for drafting the recommendations. A chair should have the ability to be objective and responsive to the viewpoints of other members. A chair whose role is facilitative must be experienced in group facilitation, have broad acceptance in the group, and have the ability to exercise authority. Although not mandatory, success is more likely if there is methodological training of the guideline panel, training for public and patient members, and professional technical support. For large, multidisciplinary panels, formal consensus developing methods have proven effective in facilitating balanced input and agreement on the final recommendations. GDGs should ensure that their objectives are commensurate with available resources and time.

***Author Disclosures***: R.K. and A.F. reported no commercial interests relevant to subject matter. F.C. consulted with and received lecture fees from NICE (National Institute for Clinical Excellence, U.K.). T.J.W., A.Q., M.L., and M.K. reported no commercial interests relevant to subject matter. G.G. and H.J.S. were members of the GRADE Working Group, which received honoraria and lecture fees regarding GRADE. All workshop participants disclosed receipt of honoraria and travel reimbursement from the American Thoracic Society.

### References

1. Cluzeau F, Wedzicha JA, Kelson M, Corn J, Kunz R, Walsh J, Schünemann H. Stakeholder involvement: how to do it right. Article 9 in integrating and coordinating efforts in COPD guideline development. An official ATS workshop report. *Proc Am Thorac Soc* 2012;9: 269–273.
2. Kelson M, Akl E, Bastian H, Cluzeau F, Curtis R, Guyatt G, Montori VM, Oliver S, Schünemann HJ. Integrating values and consumer involvement in guidelines with the patient at the center. Article 8 in integrating and coordinating efforts in COPD guideline development. An official ATS workshop report. *Proc Am Thorac Soc* 2012; 9:262–268.
3. Fretheim A, Schunemann HJ, Oxman AD. Improving the use of research evidence in guideline development: 3. Group composition and consultation process. *Health Res Policy Syst* 2006;4:15.
4. Fretheim A, Schunemann HJ, Oxman AD. Improving the use of research evidence in guideline development: 5. Group processes. *Health Res Policy Syst* 2006;4:17.
5. Burgers JS, Grol R, Klazinga NS, Makela M, Zaat J. Towards evidence-based clinical practice: an international survey of 18 clinical guideline programs. *Int J Qual Health Care* 2003;15:31–45.
6. Scottish Intercollegiate Guideline Network. SIGN 50: a guideline developer's handbook (accessed 2011 June 25). 2008. Available from: http://www.sign.ac.uk/guidelines/fulltext/50/index.html
7. National Institute for Clinical Excellence. The guidelines manual (accessed 2011 June 26). Available from: http://www.nice.org.uk
8. Hill J, Bullock I, Alderson P. A summary of the methods that the national clinical guideline centre uses to produce clinical guidelines for the national institute for health and clinical excellence. *Ann Intern Med* 2011;154:752–757.
9. Qaseem A, Snow V, Owens DK, Shekelle P. The development of clinical practice guidelines and guidance statements of the American College of Physicians: summary of methods. *Ann Intern Med* 2010;153: 194–199.
10. National Health and Medical Research Council. A guide to the development, implementation and evaluation of clinical practice guidelines (accessed 2011 June 26). Available from: http://www.nhmrc.gov.au/publications/synopses/cp30syn.htm
11. Antman EM, Lau J, Kupelnick B, Mosteller F, Chalmers TC. A comparison of results of meta-analyses of randomized control trials and recommendations of clinical experts: treatments for myocardial infarction. *JAMA* 1992;268:240–248.
12. Field MJ, Lohr KN. Clinical practice guidelines: directions for a new program. Washington, DC: National Academic Press; 1990.
13. Raine R, Sanderson C, Hutchings A, Carter S, Larkin K, Black N. An experimental study of determinants of group judgments in clinical guideline development. *Lancet* 2004;364:429–437.
14. Leape LL, Weissman JS, Schneider EC, Piana RN, Gatsonis C, Epstein AM. Adherence to practice guidelines: the role of specialty society guidelines. *Am Heart J* 2003;145:19–26.
15. Vinokur A, Burnstein E, Sechrest L, Wortman PM. Group decision making by experts: field study of panels evaluating medical technologies. *J Pers Soc Psychol* 1985;49:70–84.
16. Pagliari C, Grimshaw J. Impact of group structure and process on multidisciplinary evidence-based guideline development: an observational study. *J Eval Clin Pract* 2002;8:145–153.
17. Nationale Versorgungsleitlinie. Typ-2-Diabetes (accessed 15 Nov 2012). Available from: http://www.versorgungsleitlinien.de/themen/diabetes2/index_html
18. Moynihan R, Oxman AD, Lavis J, Paulsen E. A review of organizations that support the use of research evidence in developing guidelines, technology assessments, and health policy, for the WHO Advisory Committee on Health Research. Oslo: Norwegian Knowledge Centre for the Health Services; 2006.
19. Eccles M, Mason J. How to develop cost-conscious guidelines. *Health Technol Assess* 2001;5:1–69.
20. Murphy MK, Black NA, Lamping DL, McKee CM, Sanderson CF, Askham J, Marteau T. Consensus development methods and their use in clinical guideline development. *Health Technol Assess* 1998;2: i-88.
21. German disease management guidelines (accessed 2007 Sep 25). Available from: http://www.versorgungsleitlinien.de/english
22. Coyne DW. Influence of industry on renal guideline development. *Clin J Am Soc Nephrol* 2007;2:3–7.
23. Boyd EA, Akl EA, Baumann M, Curtis JR, Field MJ, Jaeschke R, Osborne M, Schünemann HJ. Guideline funding and conflicts of interest. Article 4 in integrating and coordinating efforts in COPD guideline development. An official ATS workshop report. *Proc Am Thorac Soc* 2012;9:234–242.

24. Scottish Intercollegiate Guidelines Network (SIGN). Guideline development process: SIGN methodology (accessed 14 Nov 2012). Available from: http://www.sign.ac.uk/methodology/index.html
25. World Health Organization. Rapid advice guidelines on pharmacological management of humans infected with avian influenza A (H5N1) virus (accessed2006 Oct. 18). Available from: http://www.who.int/csr/disease/avian_influenza/guidelines/pharmamanagement/en/index.html
26. Elwyn G, Greenhalgh T, Macfarlane F, Koppel S. Groups: a guide to small group work in healthcare, management, education and research. Radcliffe Publishing Ltd.; 2000.
27. Moreira T, May C, Mason J, Eccles M. A new method of analysis enabled a better understanding of clinical practice guideline development processes. *J Clin Epidemiol* 2006;59:1199–1206.
28. Kelson M. The NICE Patient Involvement Unit. Evidence-based Healthcare &. *Public Health* 2005;9:304–307.
29. National Collaborating Centre for Primary Care. Chronic fatigue syndrome/myalgic encephalomyelitis (or encephalopathy): diagnosis and management of CFS/ME in adults and children. NICE clinical guideline 53 (accessed 2007 Sep 25). Available from: http://guidance.nice.org.uk/CG53/niceguidance/pdf/English

# Guideline Funding and Conflicts of Interest

## Article 4 in Integrating and Coordinating Efforts in COPD Guideline Development. An Official ATS/ERS Workshop Report

**Elizabeth A. Boyd, Elie A. Akl, Michael Baumann, J. Randall Curtis, Marilyn J. Field, Roman Jaeschke, Molly Osborne, and Holger J. Schünemann; on behalf of the ATS/ERS Ad Hoc Committee on Integrating and Coordinating Efforts in COPD Guideline Development**

***Introduction*: Professional societies, like many other organizations around the world, have recognized the need to use more rigorous processes to ensure that healthcare recommendations are informed by the best available research evidence. This is the fourth of a series of 14 articles prepared to advise guideline developers in respiratory and other disease. It focuses on commercial funding of guidelines and managing conflict of interest effectively in the context of guidelines.**

***Methods*: In this review, we addressed the following topics and questions. (*1*) How are clinical practice guidelines funded? (*2*) What are the risks associated with commercial sponsorship of guidelines? (*3*) What relationships should guideline committee members be required to disclose? (*4*) What is the most efficient way to obtain complete and accurate disclosures? (*5*) How should disclosures be publicly shared? (*6*) When do relationships require management? (*7*) How should individual conflicts of interest be managed? (*8*) How could conflict of interest policies be enforced? The literature review included a search of PubMed and other databases for existing systematic reviews and relevant methodological research. Our conclusions are based on available evidence, consideration of what guideline developers are doing, and workshop discussions.**

***Results and Discussion*: Professional societies often depend on industry funding to support clinical practice guideline development. In addition, members of guideline committees frequently have financial relationships with commercial entities, are invested in their intellectual work, or have conflicts related to clinical revenue streams. No systematic reviews or other rigorous evidence regarding best practices for funding models, disclosure mechanisms, management strategies, or enforcement presently exist, but the panel drew several conclusions that could improve transparency and process.**

## INTRODUCTION

Professional societies, like many other organizations around the world, have recognized the need to use more rigorous processes to ensure that health care recommendations are informed by the best available research evidence.

Clinical practice guidelines increasingly inform clinician and patient decisions about appropriate medical care (1–4). They also carry weight in malpractice litigation and influence reimbursement policies by third-party payers (5). Ideally based on rigorous, systematic evaluations of high-quality evidence, clinical practice guidelines are intended to provide clinicians with impartial, unbiased treatment recommendations (6). Because of their direct impact on clinical practice, guidelines must be insulated from outside biases and competing interests (1).

There are, however, a number of potential threats to the impartiality and perceived independence of developers of clinical practice guidelines. This review focuses on two possible sources of bias in clinical practice guidelines: commercial sponsorship of guideline development and conflicts of interest among guideline committee members. The review examines the available evidence and current practices of select professional societies and other groups developing practice guidelines. As we will describe below, it is the fourth of a series of 14 articles prepared to advise guideline developers in respiratory and other disease on different issues.

## COMMERCIAL SPONSORSHIP

The development, maintenance, and revision of clinical practice guidelines is a costly, labor-intensive endeavor (7). Many professional societies and other groups developing guidelines rely, at least in part, on commercial sponsors to cover some of these costs. However, as noted by the Cochrane Collaboration Steering Group (8), the perception that a for-profit commercial entity (e.g., pharmaceutical and medical device companies) influenced the conclusions and recommendations of a clinical practice guideline committee could undermine the credibility of both the guidelines and the group that produced it (6, 9). Of particular concern is the possibility that guideline developers will feel—or be perceived to be—beholden to or pressured by the commercial sponsor to make recommendations favorable to the sponsor's interests.

## CONFLICTS OF INTEREST

A conflict of interest exists when an individual's personal interests (e.g., direct and indirect financial or intellectual) have the potential to compete with or influence behavior related to the individual's professional interests or obligations (i.e., evaluating the evidence and drafting recommendations for clinical practice guidelines). Biases resulting from conflicts of interest may be conscious or unconscious (10) and may influence choices made throughout the guideline development process, including conceptualization of the question, choice of comparisons, interpretation of the evidence, and, especially, drafting of the recommendations (11). Regardless of its source and its type, bias associated with conflicts of interest may damage the medical profession's trust in the integrity of the guidelines and the public's trust in science (12). Recent editorials have called for professional medical organizations to reject all industry funding for practice guidelines or outcome measures (5) and to hold guidelines panelists to the most stringent conflict-of-interest standards (13). An Institute of Medicine (IOM) report has called for adequate firewalls between funders and those developing guidelines (14) and for complete transparency in the process of guideline development (15).

---

This article is a section of "Integrating and Coordinating Efforts in Chronic Obstructive Pulmonary Disease (COPD) Guideline Development," an American Thoracic Society (ATS) and European Respiratory Society (ERS) Workshop Report. This official ATS/ERS Workshop Report was adopted by the ATS Board of Directors, August 2012, and by the ERS Executive Committee, February 2012.

Funding for this conference was made possible (in part) by 1R13HL 90485–01 from the National Institutes of Health. The views expressed in written conference materials or publications and by speakers and moderators do not necessarily reflect the official policies of the Department of Health and Human Services; nor does mention of trade names, commercial practices, or organizations imply endorsement by the U.S. Government.

**Proc Am Thorac Soc Vol 9, Iss. 5, pp 234–242, Dec 15, 2012**

**DOI: 10.1513/pats.201208-057ST**
**Internet address: www.atsjournals.org**

Financial conflict of interest is the most well-known type of conflict. This situation involves individual guideline committee members who have personal financial interests in companies developing or marketing pharmaceutical products under review. These personal financial interests include employment, consultancies, paid expert testimony, stock holdings, endowments, patents, royalties, honoraria, and in-kind gifts (e.g., travel, accommodation, meals, frequent flier miles). Intellectual conflict of interest is increasingly recognized as a source of potential bias (16) resulting from "academic activities that create the potential for an attachment to a specific point of view that could unduly affect judgment" (17). Other types of conflict of interest include academic conflicts arising from the competition for research funding as well as conflicts related to clinical revenue streams (e.g., from performing an advanced diagnostic procedure that is under consideration for a recommendation) (17, 18). Medical specialties may use practice guidelines to enlarge an area of expertise in a competitive market (19), and competing societies may use the guideline process in ways that ultimately constrain consumer choices by suppressing competing views of best practice and alternative therapies (20).

A growing number of professional societies and other groups developing practice guidelines are formulating policies to regulate commercial sponsorship and individual conflicts of interest. The intent of these policies is to address the dilemma of utilizing the experience and insight of (possibly conflicted) experts by increasing transparency and restricting or managing relationships that potentially threaten the integrity of the guidelines and the process that produces them. There is, however, considerable variation across policies, especially those involving the management of external funding for guidelines, the disclosure process, and processes for evaluating and managing the different types of conflicts of interest of guideline committee members.

In June 2007 the American Thoracic Society (ATS) and the European Respiratory Society (ERS) convened an international workshop of methodologists and researchers from around the world to coordinate efforts in guideline development for chronic obstructive pulmonary disease (COPD) and other respiratory diseases (21). Participants of the workshop completed the work over the ensuing 4 years to produce this comprehensive workshop report. This is the fourth of a series of 14 articles prepared to advise guideline developers in respiratory and other disease on different issues, in this case commercial funding of guidelines and managing conflict of interest effectively in the context of respiratory disease guidelines.

## METHODS

This review addresses eight key topics regarding the funding of guidelines and conflicts of interest (Table 1). The lead author of this section searched the PubMed database for original qualitative and quantitative research (using commercial sponsorship, clinical practice guidelines, conflicts of interest, and disclosure) and the Cochrane Methodology Register (using conflict of interest). This search was updated for the period 2008 to 2011 following the workshop. Before the workshop, ATS staff conducted a selective review of professional societies' conflict of interest policies (April 2, 2007). The authors did not conduct systematic reviews. References from the Council of Medical Editors meeting on disclosure (Sept 2004) were consulted, as were the published policies of several organizations, including the Cochrane Collaboration, National Institute for Clinical Excellence (NICE), the U.S. Food and Drug Administration (FDA), U.S. Agency for Health Care Research and Quality (AHRQ), and the Drug and Therapeutics Bulletin. The conclusions are based on available evidence, consideration of what guideline developers are doing, and ATS/ERS workshop discussions. The conclusions in this report have then been adapted for consistency with the ATS/ERS COI policies.

**TABLE 1. QUESTIONS ADDRESSED REGARDING FUNDING OF GUIDELINES AND DEALING WITH CONFLICT OF INTEREST**

1. Commercial Sponsorship:
   - How are clinical practice guidelines funded?
   - What are the risks associated with commercial sponsorship of guidelines?
   - How should commercial and other sponsorship be made public?
2. Individual Conflicts of Interest:
   - What relationships should guideline committee members be required to disclose?
   - What is the most efficient way to obtain complete and accurate disclosures?
   - How should disclosures be publicly shared?
   - When does a relationship require management?
   - How should individual conflicts of interest be managed?
   - How could conflict of interest policies be enforced?

Reprinted from Reference 65.

## RESULTS

Database searches did not yield any systematic reviews of guideline funding practices, conflict of interest policies, or financial disclosure policies related to guideline development. Nor did we identify any empirical studies of intellectual conflict of interest, although this topic has recently been the subject of a number of commentaries and recommendations (17–19, 22). Several systematic reviews of literature found an association between commercial sponsorship of research and outcomes favorable to the sponsor and the financial ties of investigators and favorable outcomes (23–25). A number of empirical studies, case studies, and commentaries demonstrate particular aspects of industry involvement in guideline development (5, 13, 16). These systematic reviews and studies are discussed below.

### How Are Clinical Practice Guidelines Funded?

No systematic reviews investigating the funding of clinical practice guidelines were found; thus, it is not possible to determine the actual range or prevalence of funding models in use. According to Steinbrook (26), medical specialty societies are the most common sponsors of clinical practice guidelines. Independent groups, such as NICE, the AHRQ, the Cochrane Collaboration, and the World Health Organization (WHO), also produce clinical practice guidelines or evaluations of evidence using several funding mechanisms.

While some medical specialty groups or professional societies continue to accept direct corporate sponsorship for practice guideline development (3), others prohibit outright sponsorship, such as the ATS and the ERS (27, 28). Since 2004, the Cochrane Collaboration has prohibited any commercial source (including all for-profit manufacturers and providers of health care or any other for-profit source with a real or potential vested interested in the findings of a specific review) from sponsoring any Cochrane reviews, Cochrane Review Groups, and Cochrane Consumer Networks (8). Likewise, the Emergency Care Research Institute (ECRI) "accepts no grants, gifts, finder's fees or consulting projects from … medical device or pharmaceutical firms." (29) In addition, the ECRI Institute prohibits any advertising from the medical device and pharmaceutical industries, and does not permit the use of its name or studies in advertising or for promotional purposes. It is important to note that the ECRI Institute, as a nonprofit entity, depends on "subscriptions" from hospitals using its assessments of products or technologies for funding support. Some of these hospitals may be for-profit and thus "commercial" (29). The Drug and Therapeutics Bulletin,

published by the British Medical Journal (BMJ) Publishing Group, states that it is "wholly independent of Government, regulatory authorities, industry and advertising, or any other form of commercial sponsorship" (28).

Other groups adopt a somewhat less stringent approach to commercial sponsorship, allowing financial support, but requiring that it either come from entities unrelated to the topic of the review (i.e., banking, airlines) or that any funding from health-related entities be placed in a general fund to be managed by the organization. This practice may help insulate specific guideline committees from direct influence and build a "firewall" between corporate funds and the work of the committee members, but may also leave the organization vulnerable to institutional conflicts of interest. The WHO accepts funding from commercial entities whose activities are unrelated to health and whose products are not harmful to health (i.e., excluding entities related to tobacco, firearms, and alcohol) (30). Although the Cochrane Collaboration prohibits all commercial funding for any review, group, or network, it will continue to allow commercial support for "non-direct, non-core" Centre activities, including translation of reviews (2006 amendment) (8). The American College of Chest Physicians (ACCP) accepts industry funding, but requires that all support be unrestricted and independent of the guideline process. Furthermore, guidelines may not have a single-source industry sponsor, industry sponsors are not revealed to members of an evidence-based guideline panel, all pharmaceutical or industry products mentioned in the guideline are referred to by generic names, and upon publication of the guideline or presentation of resulting courses, the names of the sponsors are made public (31).

### What Are the Risks Associated with Commercial Sponsorship of Guidelines?

There are significant risks associated with corporate sponsorship of clinical practice guidelines. Eichacker and coworkers in 2006 described Eli Lilly's involvement in practice guidelines for the treatment of severe sepsis and the financial interests that minimized the possible magnitude of harm associated with their product, Xigris (9). Johnson and Stricker in 2010 documented the problematic actions of the Infectious Diseases Society of America in developing clinical practice guidelines for Lyme disease and Connecticut's successful anti-trust case against the group (20).

A 2006 study of U.S. FDA Advisory Committees found that advisory committee members regularly disclose financial interests of considerable monetary value, but are rarely recused from the committees. The authors discovered a weak relationship between conflicts with competing companies and voting behavior; however, excluding advisory committee members with conflicts of interest would not have altered the vote outcome at any specific meeting (32).

The literature on the commercial sponsorship of research may also be relevant to guideline committees. As mentioned above, over the past decade, there have been a number of studies demonstrating a clear association between industry sponsorship of research and outcomes (or their interpretation) favorable to the sponsor's product (23–25, 33, 34). This relationship extends across study types and medical specialties. Unfavorable results of economic analyses of oncology drugs were less likely to be reported when the study is funded by a pharmaceutical company (35), and industry sponsorship of randomized controlled trials in psychiatry and dermatology were significantly more likely to report positive results than non–industry-funded trials (36, 37). Among the highest quality clinical research (included in Cochrane reviews), industry sponsored studies were 5.3 times more likely to support their sponsors' products than non–commercially sponsored research with similar results (34, 38). Although there are alternative ways to account for these patterns of association, these consistent findings suggest that industry involvement in research may influence the published results and raise the possibility that the evidence behind evidence-based practice guidelines—and the interpretation of these results—is, therefore, suspect (3, 39–42).

### What Information Are Guideline Committee Members Required to Disclose?

In addition to the disclosure of commercial guideline sponsors, individual members of guideline committees are usually required to disclose their personal financial relationships with related entities or sponsors. Most organizations and groups now specify a financial disclosure threshold for participants or committee members (Table 2). These range from $0 to $10,000 and involve any activity for which the member was compensated, including employment, consulting, research grants, honoraria, travel, and other reimbursements. Some groups also require disclosure of gifts, loans, intellectual property, and positions on boards and advisory boards (compensated or not). Stock ownership, including options, is also usually required to be disclosed, although some groups may not require the disclosure of mutual fund ownership. Many groups also include disclosure of nonfinancial competing interests (such as religious beliefs, participation in interest groups, or other organizations that might have an interest in the outcome of the committee's work). For example, the ATS asks guideline panel members to disclose their financial and intellectual conflicts of interest (16); the request for comprehensive disclosure of all intellectual, academic, professional, and personal conflicts that could bear upon the guideline development process has been reinforced and implemented by Guyatt and colleagues in the recent ACCP antithrombotic guidelines (17). The process required methodologists without potential conflicts to lead the formulation of recommendations in collaboration with experts who may be conflicted to a degree that would not preclude them from participation. It also required a transparent description and publication of all perceived direct financial, as well as intellectual COI.

### What Is the Most Efficient Way to Obtain Complete and Accurate Disclosures?

No randomized controlled trials or other rigorous studies evaluating different methods for obtaining conflict of interest disclosures were identified. In general, there are several formats for collecting information from committee members or participants: minimal requests to disclose any professional or financial affiliations, detailed instructions that request members to describe all involvements with organizations or entities with direct financial interest in the topic under consideration, and detailed structured checklists requiring the declaration of specific interests (43). Krimsky (43) and others are critical of the utility of minimal and open-ended requests. Bero and colleagues caution that simple disclosure requests may not reveal the nature and extent to which commercial interests exert influence over the scientific process (44). Ilakovac and coworkers warn that researchers have poor recall about the specifics of their activities, which may confound open-ended disclosure statements as an evaluation tool (45).

The Cochrane Collaboration Steering Group, the FDA advisory committee, NICE, the ATS, and the ERS use structured disclosure forms and request information on a range of financial ties, including research funding, paid consultancies, honoraria, equity holdings, gifts, patents, and royalties. The Cochrane Collaboration

**TABLE 2. COI POLICIES: SUMMARY OF SELECTED OTHER SOCIETIES**

| | ACP | ACCP | AMA | SCCM | ICBMJE | WHO |
|---|---|---|---|---|---|---|
| Definition of COI includes | | | | | | |
| Financial: | | | | | | |
| Employment | X | X | X | X | X | X |
| Consultancies | X | X | X | X | X | X |
| Stock | X | X | X (if significant) | X (major stockholder) | X | X |
| Mutual fund holdings | No | No | X (if significant) | X | N/A | No |
| Honoraria | X | X | No; given to AMA | X | X | X |
| Paid expert testimony | X | X | X | X | X | X |
| Family members | X | X | X | X | N/A | X (partner) |
| Research grants | X | X | N/A | X | X | X |
| Patents | X | N/A | X | X | N/A | X |
| Royalties | X | X | X | X | N/A | X |
| Loans | N/A | N/A | X | X | N/A | N/A |
| Gifts | N/A | N/A | X | X | N/A | N/A |
| Intellectual: pre-existent beliefs | N/A | X | N/A | N/A | X | N/A |
| Timing of disclosure | | | | | | |
| Not specified | | | | X | | X |
| Upon completion of project proposal | | X | | | | |
| Upon appointment of committee | X | X | | | | |
| Start of every committee meeting | | X | | | | |
| Completion of draft manuscript | | | | | X | |
| Completion of final pre-pub manuscript | | X | | | | |
| Other | | | X (annually) | | | |
| Level of disclosure | | | | | | |
| No amount specified | X | | X | | X | X |
| Any amount | | X | | X | | |
| >$1,000 | | | | | | |
| >$3,000 | | | | | | |
| >$5,000 | | | | | | |
| >$10,000 | | | | | | |
| Other | | | | | | |
| Direct payment | | | | | | |
| Payment to research accounts | | X | | | | |
| Payment to institution | | X | | | | |
| Disclosure period | | | | | | |
| <12 mo | | | X | | X | |
| 1 yr | X | X | | X | | |
| 3 yr | | | | | | |
| 5 yr | | | | | | |
| Other | x–4 yr | | | | | |
| COI review elements | | | | | | |
| Not specified | | | | | X | X |
| Panel chairs | X | X | X | X | | |
| Panel members | X | X | X | | | |
| Reviewers | N/A | X | N/A | | | |
| Is there a dedicated COI review body? | N/A | X | N/A | | | |
| Management strategies | | | | | | |
| Not specified | | | | | X | |
| Explicit exclusion criteria? | X | X | N/A | N/A | | X |
| Criteria customized to project role? | X | X | N/A | N/A | | |
| Recusal from discussion | X | X | X | X | | X |
| Recusal from voting | X | X | X | X | | |

This Table has been compiled by staff to the ATS Documents Development and Implementation Committee. Data reflect analysis of organizational policies available online, including those of the American College of Physicians (ACP), American College of Chest Physicians (ACCP), American Medical Association (AMA), Society for Critical Care Medicine (SCCM), International Committee of Biomedical Journal Editors (ICBMJE), and the World Health Organization (WHO). Each organization's policy was evaluated on the following: Definition of COI; Timing of Disclosure; Level of Disclosure; Disclosure Period; COI Review Elements; Management Strategies. Additional information about specific policies has been provided below.

Most of the societies provide an illustrative list of possible financial conflicts of interest rather than an exhaustive one. If one of the above listed examples was not mentioned in a policy, a "N/A" was placed in that category; however, if the policy explicitly stated that there was not a specific conflict of interest, "No" was placed in that category. Also, if a policy did not contain enough information pertaining to a specific question, "N/A" was placed in the category. X indicates that the policy does mention/require the item.

The SCCM states: "Financial interest or other relationship can include such things as… [being a] major stockholder." However, a percentage of stock ownership constituting "major" is not reported.

The AMA states: "Only those investments that constitute a significant financial investment raise a concern about a possible conflict of interest," and that " 'ownership of a material financial interest' shall mean holding a financial ownership interest of 5% or more, or holding a financial ownership interest which contributes materially to the Trustee's, Member's or Senior Manager's income, or holding a position as proprietor, director, managing partner or key employee."

The AMA requires that all honoraria received by the individual for AMA-related engagements shall be given to the association. The AMA also requires council and committee members to complete a COI form annually.

The ICBMJE requires authors to describe the role of any study sponsors, if they should exist. Also, ICBMJE states: "Editors should avoid selecting external peer reviewers with obvious potential conflict of interests, for example, those who work in the same department of institution as any of the authors."

The information from *The Endocrine Society* was reviewed, but it was not evaluated using this tool. The guidelines presented offer suggestions as to what should be done in the event of a possible conflict of interest rather than listing rules that are to be strictly enforced by a formal policy. A variety of possible COIs in areas including publications, relations with industry, clinical research, basic research, clinical practice, and the training of physicians and scientists are listed, and possible recommendations as how to rectify the situations are given.

Staff from the ACCP and the SCCM reviewed and confirmed this chart for accuracy. Individuals from the other societies were not available for comment.

Reprinted by permission from Reference 15.

also requests information on positions of management in a related entity, including service as a director, officer, partner, trustee, or employee, as well as information on outstanding loans from the entity. NICE requests information regarding an individual's private practice that could be affected by the outcome or discussion of a particular matter or product.

There is considerable variation along other dimensions of disclosure, including the timing and frequency of disclosure, the level of interest disclosed, the period of time covered by the disclosure, the amount and detail of information disclosed, and the individuals covered by the disclosure. All disclosures rely on individual self-report and there is no standardized method for verifying the accuracy or completeness of most disclosures. Historically, professional societies and other organizations have not actively pursued the accuracy or completeness of their members' financial disclosures.

### How Should Disclosures be Publicly Shared?

In general, disclosure of funding sources (and conflicts of interest) has been highly variable. The International Committee of Medical Journal Editors' Uniform Requirements for Manuscripts Submitted to Biomedical Journals (ICMJE) requires that authors, including guideline authors, disclose their affiliations, funding sources, and financial interests upon submission to the journal. Many of the 500 journals adhering to the uniform requirements now regularly publish this information as part of the manuscript (46, 47). However, the adequacy of disclosures in scientific articles has been questioned (47–49) and, even when financial sponsorship is disclosed, few studies describe the role of the sponsor (50, 51). We are not aware of any journal requiring the disclosure of other types of conflicts of interest (e.g., intellectual).

A 2000 *Lancet* study of 431 practice guidelines developed by specialty societies showed that 67% did not report the type of professionals (stakeholders) involved in the guideline development (2), and a 2005 *Nature* study of over 200 clinical practice guidelines in the U.S. National Guideline Clearinghouse showed that half of the guidelines reported no information regarding funding sources or financial conflicts of interest of the authors (52).

### When Does a Relationship Require Management?

Few organizations specify exactly which relationships constitute significant conflicts—those deemed in need of overt management or prohibited altogether. The Caring for Australians with Renal Impairment (CARI) Guidelines for Declaring Conflict of Interest prohibit members with stockholdings, employment, or other serious interests in companies active in the clinical area under consideration from participating in the activities of the guideline working group (53). The ACCP does not prohibit activities outright, but outlines a series of principles and questions to evaluate disclosed conflicts. The questions include whether that individual favors any outside entity or appears to have an incentive to do so, whether the current engagements of the individual present any conflicts between outside interests (e.g., fiduciary positions with other organizations), and whether the activity's agenda and/or content receive peer review prior to its initiation (31). The U.S. FDA does not prohibit financial relationships among its Advisory Committee members and regularly issues waivers for disclosed conflicts of interest when (*1*) the FDA determines that "the disqualifying financial interest is not so substantial that it is likely to affect the integrity of an employee's services to the government" and (*2*) the "need for the employee's services outweighs the potential conflicts of interest" (54). In making these determinations, the FDA evaluates

> the type of interest creating the disqualification; the identity of the person whose financial interest is at issue; the dollar value of the disqualifying financial interest including its value in relationship to the individual's overall assets; the nature and importance of the individual's role in the matter, including the extent to which the employee is called upon to exercise discretion; the sensitivity of the matter; and the need for the employee's services in the particular matter. (54)

The use of blanket prohibitions has been controversial and most organizations seem to prefer to retain the right to waive "disqualifying" relationships or make allowances for the expertise of the individual, the needs of the organization, and the likelihood that the financial relationships will affect the individual's judgment in a significant way.

In light of the evidence suggesting that bias induced by monetary interests is unconscious and independent of the amount of money involved (55), some organizations do prohibit certain types of relationships. The Association of American Medical Colleges (AAMC) prohibits financial relationships between principal investigators and commercial sponsors of clinical trials, but it uses a "rebuttable presumption" clause to allow the prohibition to be waived when the benefits of the research outweigh the risks of the conflict of interest (56). The U.S. National Institutes of Health and National Science Foundation establish financial thresholds for disclosure—$10,000 (recently reduced to $5,000) in annual income or 5% equity ownership in a commercial entity related to the scientific work—but do not indicate whether financial ties above these thresholds should be prohibited, avoided, or managed (57). The CARI Guidelines prohibit participation in Working Group activities for any member holding stock, receiving more than $10,000 per year in compensation for employment or other work (speaker fees, advisory fees), or "any other direct or pecuniary interest considered serious in the view of the declarer" in a pharmaceutical company active in the clinical area under consideration (53). The NIH Consensus Development Program does not allow panel participants with any financial, career, or advocacy interests in the topic under consideration. Experts with such interests may testify before the panel, but the panel members themselves are independent and without personal or financial interests (58).

### How Should Identified Conflicts of Interest Be Managed?

According to the ATS selective survey of professional societies (16), very few societies' published policies specify management strategies, describe specific exclusion criteria, or explain how criteria will be applied to individual disclosures. All mention the possibility that a member with a serious conflict of interest may be excused from the discussion. The Cochrane Collaboration specifies that "people with a direct financial interest in a particular intervention should not be involved in the review of that intervention, either as reviewers, editors, or peer reviewers" (8). The *Medical Journal of Australia* has recommended that all guideline committees have a formal process in place to assess potential conflicts of interest, that all members of the group be involved in the discussion of the management of "significant" relationships, that members with relationships need not be excluded, but the group must decide its threshold for exclusion, and that there must be complete disclosure of all relationships (4).

Public disclosure or prohibiting members from participating in discussions represent extreme ends of the possible available management strategies. The only empirical studies of management decisions in conflicts of interest detail a number of commonly used management strategies used by university conflict

of interest committees (60, 61). These possible management strategies include:

- Self declaration of conflicts of interest by potential participants
- Review of potential participants' conflicts of interest
- Disclosure of individual conflicts of interest to all guideline panel members
- Disclosure of conflicts of interest in publications and public presentations
- Reducing equity holdings below a specified amount
- Altering consulting agreements to ensure separation between consulting and guideline work
- Eliminating the financial tie (i.e., resigning from Board of Directors, canceling the consulting agreement; selling stock)
- Appointing oversight committees to review the scientific process and resulting research
- Disallowing the investigator to contribute to certain recommendations or to the guidelines as a whole
- Handling disputes in conflicts of interest resolution

Despite great variability in how institutions manage conflicts, disclosure of financial ties in all academic publications and presentations is the most frequently used management strategy (61, 62). Others have also noted the increasing use of disclosure as a way to handle financial ties of researchers (47, 62), suggesting that concerned entities consider transparency as the best way to minimize concerns about undue influence and bias among committee members (61, 62).

However, disclosure related specifically to guideline committee members has been shown in a number of studies to function rather poorly. In 2001, Papanikolaou and coworkers surveyed 191 clinical practice guidelines published in six major international journals and found that only 3.7% mentioned conflicts of interest at all; overall, only 18 authors disclosed a total of 24 potential conflicts of interest (64). Subsequently, Choudry and colleagues showed that most guideline authors have ongoing interactions with pharmaceutical companies and that a significant portion of them work as employees or consultants for these companies (1). Fifty-nine percent of the surveyed authors (47/80) reported that they had relationships with companies whose products were considered in the guidelines, and in only 2 of the 44 published guidelines were specific declarations of these relationships made (1). In 2005, *Nature* reported that more than one-third of guideline authors declared financial ties to relevant drug companies, affecting some 70% of the guideline panels surveyed (52). As mentioned above, there are no protocols in place to evaluate the accuracy and completeness of financial disclosures, and in none of the above cases was it possible to determine if or how guideline committee members' financial interests were reviewed and evaluated.

More recently, there have been a number of demands for more explicit guideline composition standards (including the inclusion of independent methodologists, epidemiologists, and policy experts) (22, 65), the addition of alternative viewpoints, the explicit notation of supporting and dissenting votes, and vetting the guidelines prior to publication through rigorous scientific peer review (20, 65). Figure 1 provides sequential steps and practical guidance for COI management.

### How Should Conflict of Interest Policies Be Enforced?

Little is known about effective or even common practices for administering and enforcing conflict of interest policies. Most organizations appear to convene a committee to review financial interest disclosures and, where deemed necessary, recommend management strategies. The U.S. FDA reviews all financial

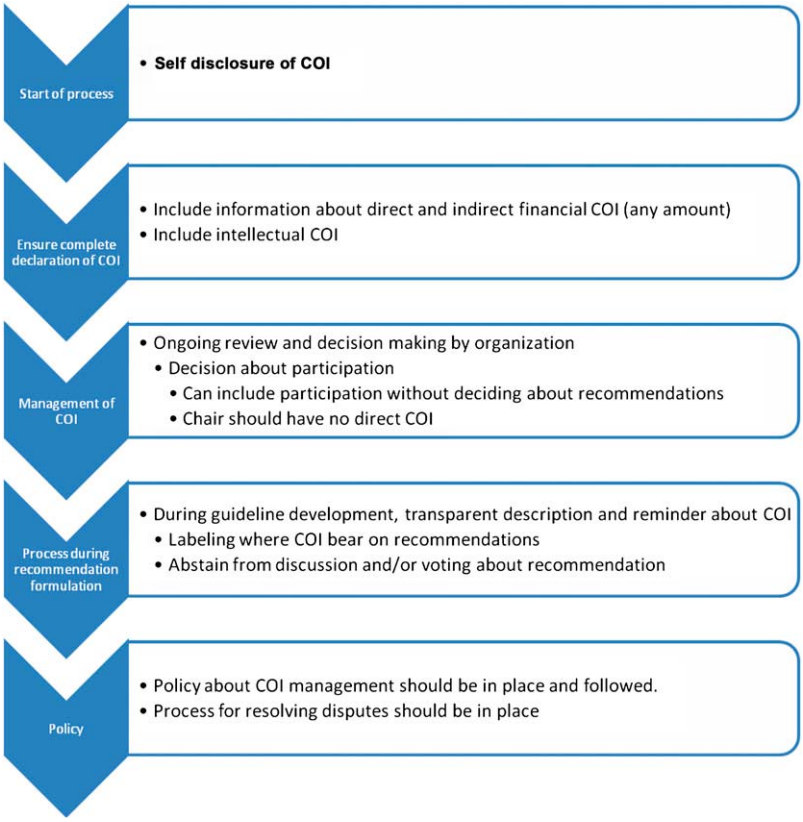


***Figure 1.*** Steps for managing conflicts of interest.

disclosure statements through a multi-stage process, beginning with initial review, followed by consultation with the individual and an FDA official, review by the FDA Ethics staff, and final approval by the appointing official. The FDA operates under federal regulations and thus has the power to enforce its decision (54).

Other organizations do not hold such broad regulatory powers; nevertheless, they may convene committees to review financial disclosure statements and make recommendations for managing conflicts or encouraging recusal of particular members. For instance, the Cochrane Collaboration directs unclear cases of financial disclosure for reviews to a "Funding Arbiter," who convenes a panel to give guidance (8).

## DISCUSSION

This review of funding and conflict of interest policies related to clinical practice guidelines raises a number of important questions. It is clear that professional societies currently depend, at least in part, on industry funding to support clinical practice guideline development. Also, many members of professional societies and guideline committees have financial relationships with commercial entities, are invested in their intellectual work, or have conflicts related to clinical revenue streams. Existing models for insulating guidelines from external bias are varied and, as yet, unstudied in any systematic way.

Public perception of the independence of clinical practice guidelines (and the committees that produce them) is critically important to their acceptance and application. Table 3 suggests a possible matrix for considering risk perception related to both commercial sponsorship and member conflicts of interest.

In attempting to balance the expert needs of guideline committees with the potential risks of close industry involvement (in both funding of guideline development and the individual members), we recommend the following.

1. How should clinical practice guidelines be funded?
   a. Groups developing guidelines should vigorously advocate for public funding for guideline development. In the absence of public funding, firewalls should be erected to insulate guideline development from the potential for—or appearance of—industry bias. Ideally, clinical practice guidelines should be developed without commercial support.
   b. Those who are funding guideline development should be free of all conflicts of interest related to the subject matter of the guideline. As an example, makers of medications to treat COPD should not fund guidelines about the management of COPD.
   c. No single source sponsorship should be allowed under any circumstances.
2. What relationships should members of guideline committees be required to disclose?
   a. Potential members of guideline committees should disclose any and all financial relationships with commercial entities whose work appears reasonably related to the topic area of the committee. This includes: any ($0 threshold) income from employment, consulting, speaking engagements, advising, and other management activities; equity ownership in public and privately traded companies; management or advisory positions (Scientific Advisory Boards, Board of Directors, Speakers Bureaus); honoraria; loans; royalties, patents, contracts, and grants; and expert testimony. Disclosures for immediate family members (spouse, partner, or child) should be made as well.
3. How should disclosures be collected and in what form?
   a. Disclosures should be made annually and discussed at the start of each committee meeting.
   b. Disclosures should be maintained on a publicly available registry (*see above*).
   c. Disclosure forms should be structured to include all disclosable relationships and be available online for members to submit.
   d. Policies should be established regarding disclosures that might include specific dollar (or equivalent) amounts; complete descriptions of consulting, research, and advisory work; and percent equity held as well as its monetary value (estimated if necessary).
4. When does a relationship constitute a significant conflict of interest?
   a. Guideline committees should strive for a balance of views represented.
   b. Policies should be developed that:
      i. Establish whether conflicts are acceptable in committee chairs. As an example, the ATS requires that at least one chair be free of conflicts.
      ii. Determine what proportion of the guideline committee is permitted to have conflicts. Currently, some organizations place no restrictions, others permit no committee members with conflicts, and others limit the conflicted members to either one-third or no more than one-half.
   c. Each guideline committee should include an independent methodologist, statistician, and policymaker to oversee and evaluate the quality of the evidence on which the guideline is based, as well as the broader health policy implications.
   d. In the event that too few individuals who are free from conflicts of interest can be located to constitute a committee, the chair must provide a clear description and documentation of the evaluation process, detailing how conflicts of interest are to be mitigated during the evaluation and recommendation process.
   e. Groups developing guidelines should elect a standing committee to oversee guideline staffing, conflict disclosures, and management of committee conflicts.

**TABLE 3. A POSSIBLE MATRIX FOR CONSIDERING RISK PERCEPTION RELATED TO BOTH COMMERCIAL SPONSORSHIP AND MEMBER CONFLICTS OF INTEREST**

| | Commercial Sponsorship? | |
|---|---|---|
| | Yes | No |
| Conflicted Members? | | |
| Yes | Perceived risk: High<br>Management: Prohibited | Perceived risk: Moderate<br>Management: Balance of views<br>Documentation of process |
| No | Perceived risk: Moderate<br>Management: Seek alternative funding | Perceived risk: Low<br>Management: None |
| | | Ideal |

f. All disqualifying relationships must be specified in advance, written and maintained on public websites.

g. The process of review must be clearly described and publicly available.

5. How should disclosures be publicly shared?

   a. Published guidelines and related web-based or print documents should clearly describe all funding sources and the methods for distributing funds to specific guideline committees.

   b. Ideally, all funding sources should be entered into an independent, publicly funded central registry to be available to all members and the public.

6. How should conflicts of interest be managed (Figure 1)?

   a. Potential members with disqualifying conflicts should be excused from the committee.

   b. The standing committee and the chair of the committee may impose restrictions on conflicted members (for instance, excusing them from final recommendations) or may request that members make efforts to mitigate their conflicts (for instance, halting consulting activities or reducing equity holdings). These policies should be prespecified and publicly available.

   c. Potential members with manageable conflicts may be asked to document their efforts to comply with the organization's recommendations. In situations in which committees have members with conflicts of interest, the chair should indicate where those conflicts might bear specifically on the recommendations. These members could then be excused from discussion and voting on the pertinent recommendations as determined by the chair(s) and other committee members.

7. How should conflict of interest policies be enforced?

   a. A standing committee should exist to review staffing decisions, aid chairs in appointing members, and serve an appeal and oversight function for guideline committees.

   b. Members who fail to disclose their financial relationships will be subject to sanction from the organizing group. The available sanctions should be specified in advance and maintained on the publicly available website.

An editorial published in the *Journal of the American Medical Association* (JAMA) called upon organizations that sponsor and promote guidelines to create joint codes for the management of conflicts of interest of members of guideline committees (13). In designing conflict of interest policies, organizations need to keep in mind that conflicted panel members are often those with the expertise in the topic and recommendations of interest. Their input into the guidelines is needed given the unique insights into clinical context and evidence they can offer. The challenge will be how to strike a balance between the competing goals of incorporating their insights and avoiding inappropriate influence of their conflicts of interest.

***Author Disclosures***: E.A.B., E.A.A., M.B., J.R.C., and M.J.F. reported no commercial interests relevant to subject matter. R.J. was a member of the GRADE Working Group. M.O. reported no commercial interests relevant to subject matter. H.J.S. was a member of the GRADE Working Group, which received honoraria and lecture fees regarding GRADE. All workshop participants disclosed receipt of honoraria and travel reimbursement from the American Thoracic Society.

## References

1. Choudhry NK, Stelfox HT, Detsky AS. Relationships between authors of clinical practice guidelines and the pharmaceutical industry. *JAMA* 2002;287:612–617.
2. Grilli R, Magrini N, Penna A, Mura G, Liberati A. Practice guidelines developed by specialty societies: the need for a critical appraisal. *Lancet* 2000;355:103–106.
3. Clinical practice guidelines and conflict of interest. *CAMJ* 2005;173:1297.
4. Van Der Weyden MB. Clinical practice guidelines: time to move the debate from the how to the who. *Med J Aust* 2002;176:304.
5. Rothman DJ, McDonald WJ, Berkowitz CD, Chimonas SC, DeAngelis CD, Hale RW, Nissen SE, Osborn JE, Scully JH, Thomson GE, *et al*. Professional medical associations and their relationships with industry: a proposal for controlling conflict of interest. *JAMA* 2009;301: 1367–1372.
6. Jackson R, Feder G. A simple, pragmatic strategy for guideline development. *BMJ* 1998;1998:427–428.
7. American Academy of Physician Assistants. Policy brief: clinical practice guidelines. 1997 (updated 1997 February 1, 2002; cited 2007 May 21). Available from: http://www.aapa.org/gandp/cpg.html.
8. Conflict of interest and commercial sponsorship. 2007 [updated 2007; cited 2007 June 4, 2007]; Available from: http://www.cochrane.org/handbook/26-declaration-interest-and-commercial-sponsorship
9. Eichacker P, Natanson C, Danner RL. Surviving sepsis: practice guidelines, marketing campaigns, and Eli Lilly. *N Engl J Med* 2006;355:1640–1642.
10. Dana J. Harm avoidance and financial conflict of interest. *J Med* 2003;1–18.
11. Bero LA. Tobacco industry manipulation of research. *Public Health Rep* (In press)
12. Friedman PJ. The impact of conflict of interest on trust in science. *Sci Eng Ethics* 2002;8:413–420.
13. Sniderman AD, Furberg CD. Why guideline-making requires reform. *JAMA* 2009;301:429–431.
14. Lo B, Field MJ, editors; Institute of Medicine. Conflict of interest in medical research, education and practice. Washington, DC: The National Academies Press; 2009.
15. Institute of Medicine. Clinical practice guidelines we can trust. Washington DC: The National Academies Press; 2011.
16. Schunemann HJ, Osborne M, Moss J, Manthous C, Wagner G, Sicilian L, Ohar J, McDermott S, Lucas L, Jaeschke R. An official American Thoracic Society policy statement: managing conflict of interest in professional societies. *Am J Respir Crit Care Med* 2009;180:564–580.
17. Guyatt G, Akl EA, Hirsch J, Kearon C, Crowther M, Gutterman D, Lewis SZ, Nathanson I, Jaeschke R, Schunemann HJ. The vexing problem of guidelines and conflict of interest: a potential solition. *Ann Intern Med* 2010;152:738–741.
18. Begolka WS, Elston DM, Beutner KR. American Academy of Dermatology evidence-based guideline development process: responding to new challenges and establishing transparency. *J Am Acad Dermatol* 2011;64:6:64e105–e112.
19. Shaneyfelt TM, Centor RM. Reassessment of clinical practice guidelines: go gently into that good night. *JAMA* 2009;301:868–869.
20. Johnson L, Stricker RB. The Infectious Diseases Society of American Lyme guidelines: a cautionary tale about the development of clinical practice guidelines. *Phil Eth and Hum Med* 2010;5:9.
21. Schunemann HJ, Woodhead M, Anzueto A, Buist S, Macnee W, Rabe KF, Heffner J. A vision statement on guideline development for respiratory disease: the example of COPD. *Lancet* 2009;373:774–779.
22. Scott IA, Guyatt GH. Clinical practice guidelines: the need for greater transparency in formulating recommendations. *Med J Aust* 2011;195: 29–33.
23. Bhandari M, Busse JW, Jackowski D, Montori VM, Schunemann H, Sprague S, Mears D, Schemitsch EH, Heels-Ansdell D, Devereaux PJ. Association between industry funding and statistically significant pro-industry findings in medical and surgical randomized trials. *CMAJ* 2004;170:477–480.
24. Ridker PM, Torres J. Reported outcomes in major cardiovascular clinical trials funded by for-profit and not-for-profit organizations: 2000–2005. *JAMA* 2006;295:2270–2274.
25. Bero L, Oostvogel F, Bacchetti P, Lee K. Factors associated with findings of published trials of drug–drug comparisons: why some statins appear more efficacious than others. *PLoS Med* 2007;4:e184.

26. Steinbrook R. Guidance for guidelines. *N Engl J Med* 2007;356:331–333.
27. American Thoracic Society. Guidelines for developing documents. 2007 (updated 2007; cited 2007 June 4, 2007). Available from: http://www.thoracic.org/sections/about-ats/assemblies/guidelines-for-developing-documents.pdf.
28. Drug and Therapeutics Bulletin: About Us. British Medical Journal; 2007 (updated 2007; cited 2007 May 25). Available from: http://dtb.bmj.com/.
29. ECRI Institute. Institute policies; 2007 (updated 2007; cited 2007 June 4, 2007). Available from: http://www.ecri.org.
30. World Health Organization. Guidelines for WHO guidelines. 2003 (updated 2003; cited 2007 June 4, 2007). Available from: http://whqlibdoc.who.int/hq/2003/EIP_GPE_EQC_2003_1.pdf.
31. Baumann M, Lewis SZ, Gutterman D. ACCP evidence-based guideline development: a successful and transparent approach addressing conflict of interest, funding, and patient-centered recommendation. *Chest* 2007;132:1015–1024.
32. Lurie P, Almeida CM, Stine N, Stine AR, Wolfe SM. Financial conflict of interest disclosure and voting patterns at Food and Drug Administration Drug Advisory Committee meetings. *JAMA* 2006;295:1921–1928.
33. Bekelman JE, Li Y, Gross CP. Scope and impact of financial conflicts of interest in biomedical research: a systematic review. *JAMA* 2003;289:454–465.
34. Lexchin J, Bero LA, Djulbegovic B, Clark O. Pharmaceutical industry sponsorship and research outcome and quality: systematic review. *BMJ* 2003;326:1167–1177.
35. Freidberg M, Saffran B, Stinson TJ, Nelson W, Bennett CL, Adams JR. Evaluation of conflict of interest in economic analyses of new drugs used in oncology. *JAMA* 1999;282:1453–1457.
36. Perlis CS, Harwood M, Perlis RH. Extent and impact of industry sponsorship conflicts of interest in dermatology research. *J Am Acad Dermatol* 2005;2005:967–971.
37. Perlis RH, Perlis CS, Wu Y, Hwang C, Joseph M, Nierenberg AA. Industry sponsorship and financial conflict of interest in the reporting of clinical trials in psychiatry. *Am J Psych* 2005;162:1957–1960.
38. Als-Nielsen B, Chen W, Gluud C, Kjaergard LL. Association of funding and conclusions in randomized drug trials: a reflection of treatment effect or adverse events? *JAMA* 2003;290:921–928.
39. Barnes DE, Bero LA. Why review articles on the health effects of passive smoking reach different conclusions. *JAMA* 1998;279:1566–1570.
40. Barnes DE, Bero LA. Scientific quality of original research articles on environmental tobacco smoke. *Tob Control* 1997;6:19–26.
41. Cho MK, Bero LA. The quality of drug studies published in symposium proceedings. *Ann Intern Med* 1996;124:485–489.
42. Abramson J, Starfield B. The effect of conflict of interest on biomedical research and clinical practice guidelines: can we trust the evidence in evidence-based medicine? *J Am Board Fam Pract* 2005;18:414–418.
43. Krimsky S. Science in the Private Interest. Lanham, Maryland: Rowan & Littlefield; 2003.
44. Bero LA, Glantz S, Hong M-K. The limits of competing interest disclosures. *Tob Control* 2005;14:118–126.
45. Ilakovac V, Fister K, Marusic M, Marusic A. Reliability of disclosure forms of authors' contributions. *CMAJ* 2007;176:41–46.
46. Campbell P. Declaration of financial interests. *Nature* 2001;412:751.
47. Van Kolfschooten F. Can you believe what you read? *Nature* 2002;416:360–363.
48. Krimsky S, Rothenberg LS, Stott P, Kyle G. Financial interests on authors in scientific journals: a principal investigator lot study of 14 publications. *Sci Eng Ethics* 1996;2:395–410.
49. Krimsky S, Rothenberg LS. Conflict of interest policies in science and medical journals: editorial practices and author disclosures. *Sci Eng Ethics* 2001;7:205–18.
50. Brennan TA, Rothman DJ, Blank L, Blumenthal D, Chimonas SC, Cohen JJ, Goldman J, Kassirer JP, Kimball H, Naughton J, *et al*. Health industry practices that create conflicts of interest: a policy proposal for academic medical centers. *JAMA* 2006;295:429–433.
51. Gross C, Gupta A, Krumholz H. Disclosure of financial competing interests in randomised controlled trials: cross sectional review. *BMJ* 2003;326:526–527.
52. Taylor R, Giles J. Cash interests taint drug advice. *Nature* 2005;437:1070–1071.
53. CARI: Caring for Australians with Renal Impairment. CARI guidelines for declaring conflict of interest; 2007 (updated 2007; accessed 2007 June 4, 2007). Available from: www.cari.org.au/Docs/CARI_Conflict_of_interest.pdf.
54. U.S. Food and Drug Administration. Draft guidance on disclosure of conflicts of interest for special government employees participating in FDA product-specific advisory committees (accessed 21 November 2011). Available from: http://www.fda.gov/OHRMS/DOCKETS/98fr/021202e.htm
55. Dana J, Loewenstein G. A social science perspective on gifts to physicians from industry. *JAMA* 2003;290:252–255.
56. Association of American Medical Colleges. Protecting subjects, preserving trust, promoting progress: guidelines for developing and implementing a policy concerning individual financial interests in human subjects research. Philadelphia; 2001. Available from: https://www.aamc.org/initiatives/coi/
57. Objectivity in research, 35810 (codified at 42 CFR 50). 1995.
58. NIH Consensus Development Program. National Institutes of Health; 2007 (updated 2007; accessed 2007 May 25, 2007). Available from: http://consensus.nih.gov/aboutCDP.htm#intro.
59. Boyd EA, Bero LA. Assessing faculty financial relationships with industry. *JAMA* 2000;284:2209–2238.
60. Boyd E, Lipton S, Bero L. Implementation of financial disclosure policies to manage conflicts of interest. *Health Aff* 2004;23:206–214.
61. Boyd EA, Cho MK, Bero LA. Financial conflict-of-interest policies in clinical research: issues for clinical investigators. *Acad Med* 2003;78:769–774.
62. Shamoo AE, Resnik DB. Responsible conduct of research. New York: Oxford University Press; 2003.
63. Papanikolaou GN, Baltogianni MS, Contopoulos-Ioannidis DG, Haidich AB, Giannakakis IA, Ioannidis JPA. Reporting conflicts of interest in guidelines of preventive and therapeutic interventions. *BMC Med Res Methodol* 2001;13.
64. Sniderman AD, Furberg CD. Why guideline-making requires reform. *JAMA* 2009;301:42951.
65. Boyd EA, Bero LA. Improving the use of research evidence in guideline development: 4. Managing conflicts of interests. *Health Res Policy Syst* 2006;4:16.

# Deciding What Type of Evidence and Outcomes to Include in Guidelines

## Article 5 in Integrating and Coordinating Efforts in COPD Guideline Development. An Official ATS/ERS Workshop Report

**Timothy J. Wilt, Gordon Guyatt, Regina Kunz, William MacNee, Milo A. Puhan, Giovanni Viegi, Mark Woodhead, Elie A. Akl, and Holger J. Schünemann; on behalf of the ATS/ERS Ad Hoc Committee on Integrating and Coordinating Efforts in COPD Guideline Development**

***Introduction*: Professional societies, like many other organizations around the world, have recognized the need to use more rigorous processes to ensure that health care recommendations are informed by the best available research evidence. This is the fifth of a series of 14 articles that were prepared by an international panel to advise guideline developers in respiratory and other diseases on approaches for guideline development. This article focuses on what type of evidence and outcomes to include in guidelines.**

***Methods*: In this review we addressed the following topics and questions. (*1*) What methods should be used to select important outcomes? (*2*) What types of outcomes should be considered? (*3*) What sources of evidence should be considered? (*4*) How should the importance of outcomes be ranked? (*5*) How to deal with surrogate outcomes. (*6*) What issues related to outcomes should be considered in the evidence review? (*7*) What quality of evidence should be used? (*8*) How to interpret the effect on outcomes. (*9*) How to incorporate outcomes related to harm. We based our responses on a PubMed literature review, prior reviews, relevant methodological research, and workshop discussions.**

***Results and Discussion*: Guideline panels should use transparent and systematic methods to select both the evidence and important outcomes, with input from groups that represent a wide range of expertise and constituencies. Outcomes should address both benefits and downsides, with consideration of the definitions, severity, and time course of the outcomes. Guideline panels should use a transparent approach to rank outcome importance recognizing that stakeholder and patient values and preferences may vary. Intermediate and surrogate outcomes are frequently reported, but their correlation with patient important outcomes may be low. A guideline panel should determine *a priori* the magnitude of effect judged clinically significant, factors that may influence outcome reporting, and whether different ways of measuring the outcomes permit the outcomes to be combined. Comprehensive identification of the evidence includes the use of multiple data sources. While randomized controlled trials (RCTs) provide the highest quality evidence, reviewers of evidence also need to consider nonrandomized studies such as case series, registries, and case-control studies if randomized trials are not available. This is particularly true for harms. The outcomes reported from RCTs may not always directly apply to clinical practice settings (i.e., they may not be generalizable).**

## INTRODUCTION

Professional societies, like many other organizations around the world, have recognized the need to use more rigorous processes to ensure that the best available research evidence informs health care recommendations. In an attempt to improve the quality of health care for many chronic conditions, including chronic obstructive pulmonary disease (COPD), clinical practice guidelines have been developed (1, 2). However, considerable variability exists in the quality and recommendations of guidelines, with some failing to adhere to established methodological standards (2). Critical to developing a high-quality guideline are decisions regarding what questions need to be addressed, which outcomes are important, and what evidence should be included.

Key clinical and research questions to be addressed in a guideline should be directly relevant to patient problems and constructed in a way that directs literature searches to relevant and precise answers. Optimally constructed clinical questions include four key elements, which are summarized as PICO (population, intervention, comparison, and outcome). When specifying the population and intervention for a clinical question, the setting (primary care clinics versus specialty outpatient clinics versus hospital settings) should be considered. When specifying the outcomes relevant to a question, the timing (i.e., short-term versus longer-term outcomes) and the duration of follow-up should be considered.

A number of factors can influence the outcomes that are included in guidelines, including the developers' specialties and potential conflicts of interest. Regarding the specialty, guidelines were typically developed by specialists in the past, even though most patients are initially managed by primary care providers who may have different perspectives. Regarding potential conflicts of interest, these may influence the type of evidence and outcomes that are included, the weighting and interpretation of the evidence, and the subsequent guideline recommendations. Conflict of interest is a common concern in guideline development. One study found that nearly 90 percent of guideline authors had some form of interaction with the pharmaceutical industry: 59% had received financial support to perform research and 38% had served as employees or consultants for a pharmaceutical company (3).

To improve quality and enhance implementation, clinical practice guidelines must be based upon high-quality systematic reviews of the available evidence. The AGREE collaboration describes the quality of clinical practice guidelines based on the extent to which the guidelines address potential biases and recommendations are internally valid, externally valid, and feasible for practice. Such recommendations require taking into account the benefits, harms, and costs of alternative management approaches, as well as the associated practical issues. Describing the quality of clinical practice guidelines also requires judgments about the methods used for developing the guidelines, the content of the final recommendations, and the factors linked to their uptake.

In June 2007 the American Thoracic Society (ATS) and the European Respiratory Society (ERS) convened an international

---

This article is a section of "Integrating and Coordinating Efforts in Chronic Obstructive Pulmonary Disease (COPD) Guideline Development," an American Thoracic Society (ATS) and European Respiratory Society (ERS) Workshop Report. This official ATS/ERS Workshop Report was adopted by the ATS Board of Directors, August 2012, and by the ERS Executive Committee, February 2012.

Funding for this conference was made possible (in part) by 1R13HL 90485-01 from the National Institutes of Health. The views expressed in written conference materials or publications and by speakers and moderators do not necessarily reflect the official policies of the Department of Health and Human Services; nor does mention of trade names, commercial practices, or organizations imply endorsement by the U.S. Government.

**Proc Am Thorac Soc Vol 9, Iss. 5, pp 243–250, Dec 15, 2012**

**DOI: 10.1513/pats.201208-058ST**
**Internet address: www.atsjournals.org**

workshop of methodologists and researchers from around the world to coordinate efforts in guideline development using COPD as a model (4). This is the fifth of a series of 14 articles that were prepared by an international panel to advise guideline developers in respiratory and other diseases on approaches for guideline development. This article focuses on what type of evidence and outcomes to include in guidelines. It complements three other papers published in this series: one about including cost considerations, one about evaluating the quality of evidence, and one about moving from evidence to recommendations (5–7).

## METHODS

This article addresses the questions described in Table 1. It is based on a review prepared for the World Health Organization (WHO) (8), the experience of an Agency for Healthcare Research and Quality (AHRQ)-Evidence-based Practice Center (which developed a report on spirometry for the diagnosis and management of COPD and a background paper for the American College of Physicians [ACP] Clinical Effectiveness Advisory Panel regarding the management of stable COPD [9]), an ERS/ATS task force report on outcomes for pharmacological treatment of COPD, and discussion at the workshop in 2007 (10). It is also based on a PubMed literature review, prior reviews, relevant methodological research, and workshop discussions.

We updated our search through April 2011 to identify high-priority publications related to type of evidence to include when developing clinical practice guidelines. For this updated search, we used the following search strategy: guidelines [MeSH Major Topic] AND (outcomes OR endpoint), yielding 813 hits in the past 5 years since publication of the previous report (8).

## RESULTS

### 1. What Methods Should Be Used to Select Important Outcomes?

A research protocol with key questions facilitates the critical steps in providing an effective evidence foundation. These steps include determination of patient-important outcomes; establishing how they are defined; description of their severity, variability, and relative importance; and clarification regarding how to locate sources of evidence. An analytic framework may be helpful by identifying the important interventions and outcomes along the causal pathway. An example of an analytic framework that was used to guide an evidence report and subsequent guideline development regarding the diagnosis and management of COPD is shown in Figure 1.

Question formulation using the PICO criteria is a critical step. Questions should consider all patient-important outcomes that allow evaluation of whether a net benefit from the intervention can be achieved (11). Since decisions in healthcare always come with harms or burdens, appropriate consideration of all benefits and harms is necessary to ensure that no net harm is done (12). Information regarding epidemiology and/or pathophysiology is typically incorporated as background information, but it can also be useful to focus questions on the populations of greatest interest or to provide rationale as to why certain interventions have a biologic basis for effectiveness.

**TABLE 1. DECIDING WHAT TYPES OF EVIDENCE AND OUTCOMES TO INCLUDE IN GUIDELINES**

| |
|---|
| 1. What methods should be used to select important outcomes? |
| 2. What types of outcomes should be considered? |
| 3. What sources of evidence to consider? |
| 4. How should the importance of outcomes be ranked? |
| 5. How to deal with surrogate outcomes |
| 6. What issues related to outcomes to consider in the evidence review? |
| 7. What quality of evidence should be used? |
| 8. How to interpret the effect on outcomes |
| 9. How to incorporate harms outcomes |

The systematic and explicit development of questions and identification of outcomes benefit from the input of groups representing a wide range of expertise and constituencies (13). This diversity helps ensure that the key outcomes are included, understandably derived, and directly relevant for dissemination and implementation. It may also focus the resources required to conduct evidence syntheses. For example, for the AHRQ and ACP review on the management of stable COPD, there was a technical expert panel that included individuals with expertise in the diagnosis and management of COPD, primary care, disease prevention and practice, pharmacology, policy decision making, and methodology. Representatives from various societies with wide-ranging interests and experiences in treating patients with COPD provided input. Patient representatives were not included in this particular project, but can be valuable partners in formulating key questions and determining outcomes of interest.

### 2. What Types of Outcomes Should Be Considered?

The clinical outcome(s) of interest are generally those outcomes that measure what one expects the intervention to improve or affect and are critical or important for decision making. There should be both beneficial outcomes and harmful outcomes.

As an example, for questions about treatment (e.g., inhaled bronchodilators or pulmonary rehabilitation), the outcomes of interest may be symptoms, functional status, or survival, as well as harm, resource utilization, or cost. As another example, for questions about diagnosis or prognosis (e.g., spirometry), the outcomes of interest may be diagnostic accuracy or prognostic stratification. Guidelines addressing diagnostic or prognostic questions (e.g., should spirometry be performed in asymptomatic individuals to provide them with a diagnosis of airflow obstruction or to determine prognosis in symptomatic individuals?) often require unique considerations related to the literature search, evidence assessment, data analysis, and judgments about the link between test accuracy and patient outcomes.

A detailed understanding of the definitions, severity, and time course of relevant outcomes is important. As an example, multiple definitions of COPD exacerbation have been used (e.g., the proportion having at least one COPD exacerbation, rate per 100 patient-years, exacerbations requiring hospitalization, exacerbations receiving antibiotics and oral corticosteroids, etc.). Some outcomes may be adequately assessed by small or short-term studies (e.g., change in dyspnea or development of dry mouth), but other outcomes require large or long-term studies (e.g., mortality or risk of fracture with inhaled corticosteroids). Depending upon the outcomes of interest, guideline panels need to make decisions as to the types of studies to include (i.e., short-term trials or longer-term trials), whether it is feasible and appropriate to combine findings, and what constitutes clinically relevant differences in a particular outcome. Guideline groups may also wish to include feasibility and outcomes not directly based on clinical benefits and harms, such as resource utilization.

The WHO and many other guideline panels have used an approach promulgated by the GRADE Working Group that asks guideline panels to make judgments about the importance of outcomes for decision making using rating scales (14–16). Typically, panel members follow written or verbal instructions using a nine-point rating scale to identify outcomes as not-important (score 1–3), important (score 4–6), or critical (score 7–9). This rating serves many purposes, including the reduction of work on unnecessary outcomes that may be reported in the literature but

**TABLE 2. RECOMMENDATIONS FOR DEFINING OUTCOMES AND DETERMINING WHETHER THEY CAN BE COMBINED**

| | |
|---|---|
| Outcome definitions | Specify whether outcome definitions or the way outcomes were measured differed among studies. Specify whether surrogate outcomes or combined endpoints were used. If observational studies are included, specify the definition and measurement of confounding factors/effect modifiers. |
| Primary vs. secondary outcomes | Specify whether outcomes were primary or secondary outcomes in the original studies. Specify benefit and harm outcomes and their combinations. |
| Outcome assessment in randomized controlled trials | Specify whether intention to treat, per protocol, last observation carried forward, etc. outcome definitions are used in combined randomized controlled trial analyses. If estimates from different outcome definitions were combined, then subgroup and/or sensitivity analyses should also be undertaken. |

not important or critical for decision making. Establishing the ground rules for which outcomes are considered for decision making early in the process of guideline development avoids unwanted debate and helps systematize the assessment of evidence related to these outcomes.

### 3. What Sources of Evidence Should Be Considered?

Once the key interventions and potential outcomes (harms and benefits) are identified, a detailed search strategy and study inclusion/exclusion criteria need to be constructed to identify the best sources of information. This is not always straightforward. Although driven by key issues regarding science, pragmatism also plays a role. Randomized controlled trials (RCTs) remain the gold standard for assessing the effectiveness of key interventions. Their conduct and completion for all key management questions is not always be feasible or available. Evidence synthesizers need to make a decision about whether to include other sources of evidence, such as observational studies. They also need to decide whether to include evidence from non–peer-reviewed literature. The use of gray literature (i.e., evidence found in sources other than traditional peer-reviewed printed manuscript, such as technical reports, meeting abstracts, doctoral dissertations, and websites [Food and Drug Administration {FDA}, company]) may be necessary for interventions involving emerging technologies, complementary medicine, and health systems interventions. Product labels may also be a useful source of harms data.

### 4. How Should the Importance of Outcomes Be Ranked?

Although further work is needed, relative ranking scales resemble a pragmatic and feasible approach to identifying the importance of outcomes. The primary focus when determining the relative importance of outcomes should be the person who primarily benefits from wise decision making—namely, the patient. Otherwise, outcome importance may vary according to whether it is determined by a patient, family member, clinician, researcher, health system, insurance or pharmaceutical/device industry, or society (8). A transparent approach to describing the potential outcomes (benefits and harms) helps target the outcomes that are most important to include in the evidence summary and guideline statement.

Intermediate outcomes may be useful surrogates (e.g., change in $FEV_1$). However, they should not be used instead of direct patient-centered outcomes of effectiveness (e.g., exacerbations, hospitalizations, and mortality) or harms (e.g., dry mouth, osteoporosis, and cardiovascular toxicity) unless they have been validated.

Information related to values and preferences should be incorporated if possible, but may vary depending upon the stakeholders and even among individual patients. Values are not right or wrong, and may vary according to many factors associated with different stakeholders (e.g., some patients may place greater value on symptom relief, others may prefer life prolongation, and still others may place higher priority on minimization of harms including costs).

Different methods of presenting the findings may also influence how outcomes are perceived (e.g., presenting data as mortality versus survival can alter the perception of treatment effectiveness, percentages versus rates versus number per 100, relative versus absolute risks) (17).

### 5. How to Deal with Surrogate Outcomes

For the management of stable COPD, specific treatment options include: long-acting inhaled bronchodilators used as monotherapy or in combination with inhaled corticosteroids, pulmonary rehabilitation, assisted ventilation, disease management, and supplemental oxygen therapy. Patient-important outcomes include (but are not limited to) exacerbations, health-related quality of life, hospitalizations, and mortality. Adverse events include withdrawals, treatment-specific withdrawals, all adverse events, and treatment-related adverse events. Feasibility could be assessed by compliance with therapies or expert opinion on how consistent patient populations, trial design, and trial conduct are to routine clinical practice. Surrogate measures are also frequently reported outcomes.

Surrogate outcomes are usually easy to measure and frequently require smaller sample sizes or shorter study durations to identify treatment effectiveness. Surrogate measures should represent pathophysiological processes and, by definition, are correlated with clinical outcomes. However, the certainty and degree of correlation is frequently difficult to establish. For example, in COPD trials, spirometric parameters are often used as surrogate outcome measures. These surrogate markers include response to bronchodilators or change over time in spirometry. Many trials claim effectiveness of treatments based on spirometric improvements without adequate assessment of clinical outcomes. However, their correlation with clinically relevant effectiveness outcomes is often low, as well as their ability to be useful for guiding management decisions in individual patients. Frameworks such as the one developed by Lassere and colleagues are helpful for assessing the validity and usefulness of surrogate endpoints (18).

### 6. What Issues Related to Outcomes Should Be Considered in the Evidence Review?

Defining outcomes can be difficult. It is often not feasible or appropriate to extract all recorded outcomes, and some decisions may not become apparent until after data extraction has begun (e.g., deciding whether to extract and combine different definitions of exacerbations). However, definitions of the outcomes of interest and the relative values placed on these outcomes should be done prior to data extraction and analysis. Transparency about how *a priori* decisions were made and subsequently revised is critical to having a high-quality evidence report that minimizes bias.

Development of a research protocol outlining the review and guideline process is important. The *objectives* section states the primary objective of the review, including the intervention(s) evaluated and the targeted condition, patients, settings, and

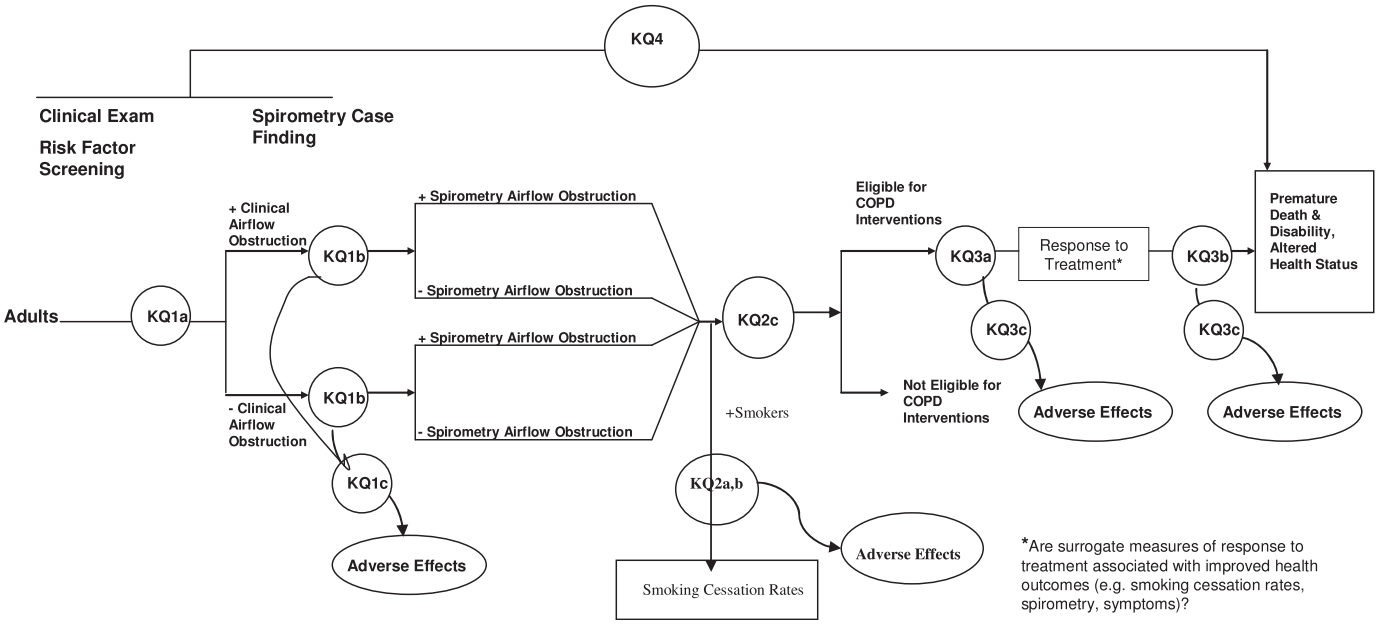


***Figure 1.*** Spirometry for case finding of COPD—analytic framework.

- (KQ1a and b) What is the prevalence of airflow obstruction as defined by (a) clinical examination or (b) spirometry in various adult populations?
- (KQ1c) What are the harms of providing a diagnosis of airway obstruction by spirometry?
  - Patient population: adults at risk for COPD
  - Intervention: clinical examination
  - Comparison: spirometry
  - Outcomes: (a) Diagnostic accuracy of components of the clinical examination for predicting presence and severity of airflow obstruction. (b) Prevalence of spirometric levels of airflow obstruction according to age, smoking status, sex, race/ethnicity, respiratory symptom status, and previous clinical diagnosis. (c) Harms: health costs and personnel, patient anxiety, unnecessary or ineffective treatment, missed opportunities to diagnosis and treat other conditions.
- (KQ2a and KQ2b) Can spirometry increase smoking cessation rates and how does patient knowledge of the spirometry outcome affect smoking cessation rates?
- (KQ2c) Can use of initial or follow-up spirometry increase the probability of initiation of successful smoking cessation treatment compared with clinical examination?
  - Patient population: adults with a current or past history of smoking
  - Intervention: spirometry
  - Comparison: usual care
  - Outcomes: Long-term biologically confirmed abstinence, patient reported abstinence according to spirometric and smoking status; incidence/prevalence of smoking cessation programs/medications. Harms: decreased abstinence, health costs and personnel effort
- (KQ3a and b) Does effectiveness of treatment vary based on (a) baseline severity or (b) change in spirometry (short term due to initial therapy or progression over time) or symptom status?
- (KQ3c) What are the harms associated with treatment based on severity or change in spirometry?
  - Patient population: adults with COPD according to spirometric and symptom status
  - Intervention: COPD-specific interventions
  - Comparison: placebo, active comparators
  - Outcomes: Effectiveness outcomes: exacerbations, validated respiratory health status measures, hospitalizations, mortality according to baseline spirometric values, short-term spirometric response to therapy, and rate of spirometric decline over time as well as symptom status. Harms: withdrawals, adverse effects (overall and specific), serious adverse effects (overall and specific).
- (KQ4) Is prognosis based on spirometry more accurate than prognosis based on clinical examination alone?
  - Patient population: adults with COPD according to spirometric classification
  - Intervention: spirometric classification
  - Comparison: clinical examination and history
  - Outcomes: Relative increase in predictive value using spirometry in addition to clinical examination for determining future risk of all-cause and disease specific mortality, hospitalizations, exacerbations

problem. The *methods* section outlines the types of studies, participants, interventions, and outcome measures. It also describes the databases to be searched (e.g., limit to PubMed and the Cochrane Library, or include Embase and the gray literature such as conference proceedings, the National Technical Information Service, FDA information, relevant industry websites, etc.), the types of studies that will be included (e.g., RCTs), the types of studies that will be excluded (e.g., case series), and the reasons for the exclusion. Any restrictions or exclusions of specific population characteristics, interventions, outcomes, or settings are noted along with the rationale. Relevant subgroups, assessment of the methodological quality of the included studies (based on allocation concealment, blinding, intention-to-treat analysis, and attrition), and how the data will be combined (if possible and appropriate) are described.

Heterogeneity between the included studies is assessed and the reasons for any heterogeneity are discussed. Sensitivity analyses are conducted to determine how changes in the definitions of outcomes may alter conclusions. For example, outcomes related to exacerbations are defined in several ways. This is related to variation in the definition of exacerbations, the definitions of severity, or the metric used to report exacerbation (e.g., percentage of subjects having at least one exacerbation, rates of exacerbations over time, exacerbations requiring hospitalizations,

exacerbations not further defined, exacerbations requiring use of antibiotics and oral corticosteroids). Evidence synthesizers need to understand the nuances associated with potential clinical heterogeneity in outcome definitions because these can play important roles in determining whether different studies can be combined, what relative emphasis to place on different outcomes or studies, and ultimately whether an intervention is deemed effective.

The AHRQ/ACP report for the recent ACP/ATS/ERS/ACCP COPD guidelines used the following outcomes of effectiveness: percentage of subjects experiencing at least one exacerbation (and included the authors' definition of exacerbation), mean change in respiratory health–related quality of life, overall and respiratory specific hospitalizations, and overall mortality. The report included other outcomes that are considered surrogates for functional status or quality of life, such as standardized walking distance in 6 minutes. The report limited analysis of surrogate outcomes, such as the spirometric change over time or the frequency of use of "rescue medications," because these are less likely to influence patient and physician treatment decision making. Respiratory health–related quality of life was assessed by the St. George Respiratory Questionnaire (SGRQ) or the Chronic Respiratory Disease Questionnaire (CRQ), validated instruments that quantify the extent of COPD impairment.

### 7. What Quality of Evidence Should Be Used?

Guidelines should be based upon the highest-quality evidence available, although the decision about what constitutes the highest-quality evidence depends on the exact nature of the question(s) and outcomes being addressed (*see* Table 2). To determine the quality of evidence, the GRADE approach recommends that the study design and eight key factors be assessed (i.e., judged transparently for their potential impact on the magnitude of effect) for each outcome. The factors include risk of bias, inconsistency, indirectness, imprecision, publication bias, magnitude of effect, dose–response relationships, and whether all plausible confounding would result in an opposite effect (Table 3) (19).

To find the evidence, a literature search strategy should be developed that states which literature sources will be searched and which studies will be selected for review. A key aspect of the search strategy is the type of study that will be selected (e.g., RCTs only or also nonrandomized studies, especially if assessing for potentially serious but infrequent harms data). Other components of the search strategy include the language (is the additional information from foreign language reports relevant to a certain guideline group; relevant to a certain clinical setting; and worth the costs, time, and inconvenience of translation and extraction?), year of publication (are early reports/patient populations so clinically outdated relative to current practice that their inclusion would be irrelevant or misleading?), study size (do small studies provide inadequate events to validly contribute summary outcome estimates or are they required to yield a comprehensive assessment of the evidence?), patient population, interventions, and comparisons of interest. For some clinical practice guideline questions, the use of gray literature is also useful.

RCTs can provide the highest quality of evidence for determining the effectiveness of interventions. However, the quality of RCTs varies greatly, and there is considerable evidence indicating that trials of lower quality are more likely to report an exaggerated effect size than studies of higher quality. Furthermore, negative trials are less likely to be published than positive studies, and methods to identify such studies and/or correct for publication bias have limitations (20). This can lead to exaggerated or biased estimates of effect size if outcomes from all conducted studies are not included.

It is important to determine whether the outcomes assessed may be influenced by key study design features. Several factors have been recommended to determine the quality of the individual RCTs, and these should be incorporated into assessing the outcomes and their magnitude of effects. The methods of Schulz are often used to assess the quality of RCTs based on allocation concealment (21). However, quality scale scores also have limitations, and they should be evaluated for usefulness on the basis of the question asked and, in particular, the outcomes being addressed. As an example, the AHRQ/ACP report for the recent ACP/ATS/ERS/ACCP COPD guidelines assessed whether outcomes vary according to study level factors, including blinding, analysis by intent-to-treat, length of follow-up, dropouts or lost to follow-up, and funding source.

The quality of the evidence can be rated according to the GRADE approach when the findings from systematic reviews and meta-analyses are incorporated into guidelines (19, 22). Assessment of the quality of evidence using GRADE is the focus of another article in this series (6).

### 8. How to Interpret the Effect on Outcomes

Merely stating what outcome(s) to include in a guideline is not sufficient. The details of the outcomes are important, including the metric used to report those outcomes (including the point estimate; confidence intervals; *P* values; and their consistency across study designs, interventions, populations, and settings). The guideline panel should determine *a priori* the magnitude of effect that is judged important to the target population (in particular in relation to effects for other outcomes), factors that may influence the reporting of outcomes, and whether different methods of outcome measurement are similar enough to allow for combining outcomes.

Research papers generally report health status or quality-of-life measures. These measures are often the mean change relative to baseline, or a comparator, and the level of statistical importance rather than patient importance. Fortunately, prior work has determined that a 4-unit reduction (out of 100) on the SGRQ and a 0.5-unit increase per question on the 7-point CRQ represent the minimal patient important difference (23–25). Studies of pulmonary rehabilitation have found that a clinically significant improvement in the distance walked during a standardized 6-minute-walk test exceeds the previously validated minimal effect size of 35 to 53 m (26, 27). Therefore, guideline developers need to evaluate whether the magnitude of effect (and associated confidence intervals) achieved levels judged to be clinically effective, rather than evaluating only the mean change and its accompanying level of statistical significance.

The metric used to measure the outcome can be important in several ways, including whether the measurements are similar enough to be combined. As an example, pooled results in one systematic review revealed no statistically significant relative or absolute reduction in the likelihood of having greater than or equal to 1 exacerbation during the trial period when combination long-acting bronchodilators and corticosteroids were compared with either agent alone (28). The longest and largest study, the TORCH trial, found a relative risk reduction of nearly identical magnitude to the pooled analysis, but the reduction was statistically significant (29). Adding the TORCH trial to the pooled analysis probably would have improved the precision of estimated effects, but it was not possible to do so because the TORCH trial reported its exacerbation outcomes as annual rates (29).

### 9. How to Incorporate Harms Outcomes

To be fully informative, the harms of an intervention should always be assessed (Table 4). RCTs are often not the best source

**TABLE 3. GRADE APPROACH TO GRADING THE QUALITY OF EVIDENCE**

| Source of Body of Evidence | Initial Rating of Quality | Factors that May Decrease the Quality | Factors that May Increase the Quality | Final Quality of a Body of Evidence |
|---|---|---|---|---|
| Randomized trials | High | 1. Risk of bias | 1. Large effect | High (⊕⊕⊕⊕) |
| | | 2. Inconsistency | 2. Dose–response | Moderate (⊕⊕⊕◯) |
| Observational studies | Low | 3. Indirectness | 3. All plausible residual confounding would reduce the demonstrated effect or would suggest a spurious effect if no effect was observed | Low (⊕⊕◯◯) |
| | | 4. Imprecision | | Very low (⊕◯◯◯) |
| | | 5. Publication bias | | |

Case reports and clinical observations generate observational study evidence that is usually downgraded because of risk of bias to very low quality.

of information to assess harms and, therefore, other sources of information are required. It is insufficient to conclude that a specific adverse event did not occur if it was not reported, since adverse events are often poorly reported in RCTs. Adjudication committees are rarely used, and these events may occur too rarely to provide conclusive answers (30). Also, considerable variability in the definitions of adverse events may exist and standardized definitions are often not provided (31), making pooling across trials neither feasible nor appropriate.

Guideline developers should consider supplementing published trial results with well-designed observational studies (i.e., case–control and population-based cohort studies), as well as unpublished data. Electronic database searches can be supplemented with other methods for identifying relevant published studies, including reviews of reference lists, soliciting experts for additional citations, and/or using the citation tracking feature on OVID. Developers may also search for unpublished trials or data from comparative effectiveness reviews (CERs) on the FDA website, Scientific Information Packets (SIPs), governmental regulatory agencies outside of the United States (e.g., the European Medicines Agency), relevant conference proceedings, and governmental or nongovernmental clinical trials databases.

Developers should pay particular attention to the methods used for ascertaining exposures and outcomes, as well as for measuring and analyzing potential confounders. They should also support the development and testing of quality ratings instruments such as the McMaster Quality Assessment Scale for Harms (McHarm) (32). Evidence on harms from each type of study should be clearly summarized in summary tables and/or in narrative format. Use of standardized definitions for reporting of harms is encouraged. Common terminology includes the following: Harms, adverse events, serious adverse events, severe adverse events, side effects, and adverse effects.

For long-acting inhaled therapies, the AHRQ/ACP evidence report assessed adverse effects, serious adverse effects, treatment adherence, study withdrawals, and withdrawals due to adverse effects in trials lasting at least 1 year in duration and in systematic reviews that specifically addressed adverse effects (9). As part of the assessment, it was determined whether the studies used placebo or active control run-in periods, as well as the number and reasons for exclusion from randomization during the run-in period of potentially eligible patients.

## SUMMARY

Clinical practice guidelines must be based on high-quality systematic reviews of the best available evidence. This improves quality and consistency, while enhancing dissemination and implementation. The types of evidence and outcomes that are used to provide that high-quality information to guideline developers depend, in part, on the nature of the question(s).

It is important to use transparency in defining and identifying the outcomes of greatest clinical interest. These outcomes should include both harms and benefits. Approaches to determining which outcomes are important or critical to decision making in guidelines include the use of simple ranking scales. The use of these scales streamlines the conduct of systematic reviews and facilitates the development of recommendations.

Key stakeholders may provide important information about evidence sources and outcomes of greatest interest. Comprehensive

**TABLE 4. SUMMARY OF RECOMMENDED APPROACHES FOR EVALUATING HARMS**

- Identify the harms that are most important to decision makers and users of the interventions being evaluated using Technical Expert Groups.
- Use consistent and precise terminology when describing evidence on harms.
- Routinely include data from randomized trials, including both head-to-head and placebo-controlled trials (especially for uncommon or rare adverse events).
- Include unpublished trials and unpublished data from published trials when publication bias or selective outcomes reporting bias are significant issues.
- Data from observational studies on harms should routinely be included, focusing on the highest-quality and most applicable observational studies. Clearly cite inclusion criteria for observational studies and rationale for excluding certain types of studies.
- Do not rely solely on electronic databases to identify studies of harms. Rather, consider searching on a broad range of sources (including government regulatory sites and clinical trials registries) and support the development of registries for observational studies of harms.
- When evaluating risk of bias (quality) of studies on harms, do not assume that assessments of harms are adequate because assessments of beneficial outcomes are appropriate. Rather, specifically evaluate how well studies define, assess, and report harms. Support the development of methods for evaluating risk of bias in studies reporting harms.
- When confidence intervals for an adverse event indicate no statistically significant risk but include the possibility of clinically significant risk, interpret the likelihood and clinical implications of the increased risk.
- Do not draw conclusions about nonequivalence or noninferiority unless there are appropriate data justifying such statements.
- Avoid assuming that class effects are present for two or more interventions unless there are clinical outcomes data supporting such assumptions. Include analyses of heterogeneity when combining data on harms from two or more interventions.
- Avoid implicit indirect comparisons. Rather, evaluate whether different sets of trials meet assumptions for similarity of treatment effects, and if so, perform formal indirect comparisons.
- Consider including data on harms from different populations when they are not likely to differ in estimates of harms. Include analyses of heterogeneity when combining data on harms from different populations.
- Do not pool data from observational studies unless there is a clear rationale to do so. When discrepancies between randomized controlled trials and observational studies are present, evaluate potential reasons for these discrepancies.
- Present data from the most important harms first, with more reliable evidence presented preceding less reliable evidence.

identification of evidence includes the use of multiple electronic data sources, in particular the Cochrane library and other sources of ongoing or completed clinical trials. While RCTs can provide the highest-quality evidence for assessing effectiveness, they may not be appropriate or adequate for evaluating harms. Identification of nonrandomized studies often requires use of the gray literature, including pharmaco-epidemiological studies.

RCTs may not report outcomes that are important to patients or generalizable to clinical practice settings. As an example, key outcomes for the management of COPD include disease-specific and overall mortality, exacerbations, hospitalization, clinical resource utilization, quality of life, compliance, and costs. Other frequently reported outcomes include measures of dyspnea, rescue medication utilization, night-time awakening, exercise tolerance, and potential physiological surrogates, such as spirometry. These are not as well validated or as important in clinical decision making.

Harms include serious adverse events, specific adverse events, and withdrawals. Unfortunately, many studies provide different definitions of these outcomes and/or report differing severities of the outcomes. Standardization of reporting of outcomes and establishing the minimal important difference is required for informed decision making.

***Author disclosures***: T.J.W. reported no commercial interests relevant to subject matter. G.G. was a member of the GRADE Working Group. R.K. reported no commercial interests relevant to subject matter. W.M. consulted with GlaxoSmithKline and was on advisory committees of Almirall, Bayer, GlaxoSmithKline and Micromet. He received lecture fees from GlaxoSmithKline, Healthcare Education Services, Healthmatter Communications and Zambon, and grants from Ceremedix, Pfizer and Unilever. He received royalties from Atlas Medical Publishing, B.C. Decker, Clinical Publishing, Healthpress, Hodder Headline, Professional Communications and Taylor & Francis, and travel awards from Astra Zeneca, Boehringer Ingelheim, GlaxoSmithKline and Micromet. M.A.P. received lecture fees from GlaxoSmithKline and Grants from Astra Zeneca, Boehringer Ingelheim and GlaxoSmithKline. G.V. consulted and was on advisory committees of Pfizer, received lecture fees from GlaxoSmithKline, and received grants from GlaxoSmithKline and Pfizer. M.W. consulted with GlaxoSmithKline and received lecture fees from Pfizer. E.A.A. reported no commercial interests relevant to subject matter. H.J.S. was a member of the GRADE Working Group, which received honoraria and lecture fees regarding GRADE. All workshop participants disclosed receipt of honoraria and travel reimbursement from the American Thoracic Society.

## References

1. Boyd CM, Darer J, Boult C, Fried LP, Boult L, Wu AW. Clinical Practice Guidelines and Quality of Care for Older Patients With Multiple Comorbid Diseases: Implications for Pay for Performance. *JAMA* 2005; 294:716–724.
2. Celli BR, MacNee W; ATS/ERS Task Force. Standards for the diagnosis and treatment of patients with COPD: a summary of the ATS/ERS position paper. *Eur Respir J* 2004;23:932–946.
3. Choudhry NK, Stelfox HT, Detsky AS, Relationships Between Authors of Clinical Practice Guidelines and the Pharmaceutical Industry. *JAMA* 2002;287:612–617.
4. Schunemann HJ, Woodhead M, Anzueto A, Buist S, Macnee W, Rabe KF, Heffner J. A vision statement on guideline development for respiratory disease: the example of COPD. *Lancet* 2009;373:774–779.
5. Hill OL Sr, Falck-Ytter Y, Cruz A, Atkins D, Baumann M, Jaeschke R, Woitalla T. Schünemann HJ, 6. Incorporating considerations of cost-effectiveness, affordability and resource implications in guideline development: An official ATS Workshop Report. *Proc Am Thorac Soc* 2010; (submitted).
6. Guyatt G, Akl E, Oxman A, Wilson K, Puhan MA, Wilt T, Gutterman D, Woodhead M, Antman E, Schünemann HJ. Synthesis, grading and presentation of evidence in guidelines. Article 7 in integrating and coordinating efforts in COPD guideline development. An official ATS/ERS Workshop Report. *Proc Am Thorac Soc* 2012;9:256–261.
7. Schunemann H, Oxman A, Akl EA, Brozek J, Montori V, Heffner J, Hill S, Woodhead M, Campos-Outcalt D, Alderson P, *et al*. Moving from evidence to developing recommendation in guidelines. Article 11 in integrating and coordinating efforts in COPD guideline development. An official ATS/ERS Workshop Report. *Proc Am Thorac Soc* 2012;9: 282–291.
8. Schunemann H, Oxman A, Fretheim A. Improving the use of research evidence in guideline development: 6. Determining which outcomes are important. *Health Res Policy Syst* 2006;4:18.
9. Qaseem A, Snow V, Shekelle P, Sherif K, Wilt TJ, Weinberger S, Owens DK. Diagnosis and management of stable chronic obstructive pulmonary disease: a clinical practice guideline from the American College of Physicians. *Ann Intern Med* 2007;147:633–638.
10. Cazzola M, MacNee W, Martinez FJ, Rabe KF, Franciosi LG, Barnes PJ, Brusasco V, Burge PS, Calverley PM, Celli BR, *et al*. Outcomes for COPD pharmacological trials: from lung function to biomarkers. *Eur Respir J* 2008;31:416–469.
11. Guyatt GH, Oxman AD, Kunz R, Atkins D, Brozek J, Vist G, Alderson P, Glasziou P, Falck-Ytter Y, Schunemann HJ. GRADE guidelines: 2. Framing the question and deciding on important outcomes. *J Clin Epidemiol* 2011;64:395–400.
12. Schunemann HJ. Guidelines 2.0: do no net harm-the future of practice guideline development in asthma and other diseases. *Curr Allergy Asthma Rep* 2011;11:261–268.
13. Kunz R, Fretheim A, Cluzeau F, Wilt TJ, Qaseem A, Lelgemann M, Kelson M, Guyatt G, Schünemann HJ, Guideline group composition and group processes. Article 3 in integrating and coordinating effeorts in COPD guideline development. An official ATS workshop report. *Proc Am Thorac Soc* 2012;9:229–233.
14. Falzon D, Jaramillo E, Schunemann HJ, Arentz M, Bauer M, Bayona J, Blanc L, Caminero JA, Daley CL, Duncombe C, *et al*. WHO guidelines for the programmatic management of drug-resistant tuberculosis: 2011 update. *Eur Respir J* 2011;38:516–528.
15. Hsu J, Brozek JL, Terracciano L, Kreis J, Compalati E, Stein AT, Fiocchi A, Schunemann HJ. Application of GRADE: Making evidence-based recommendations about diagnostic tests in clinical practice guidelines. *Implement Sci* 2011;6:62.
16. Fiocchi A, Schunemann HJ, Brozek J, Restani P, Beyer K, Troncone R, Martelli A, Terracciano L, Bahna SL, Rance F, *et al*. Diagnosis and rationale for action against cow's milk allergy (DRACMA): a summary report. *J Allergy Clin Immunol* 2010;126:1119–1128.
17. Zikmund-Fisher BJ, Fagerlin A, Ubel PA. Mortality versus survival graphs: improving temporal consistency in perceptions of treatment effectiveness. *Patient Educ Couns* 2007;66:100–107.
18. Lassere MN, Johnson KR, Boers M, Tugwell P, Brooks P, Simon L, Strand V, Conaghan PG, Ostergaard M, Maksymowych WP, *et al*. Definitions and validation criteria for biomarkers and surrogate endpoints: development and testing of a quantitative hierarchical levels of evidence schema. *J Rheumatol* 2007;34:607–615.
19. Guyatt GH, Oxman AD, Vist GE, Kunz R, Falck-Ytter Y, Alonso-Coello P, Schunemann HJ, Group GW. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ* 2008; 336:924–926.
20. Lau J, Ioannidis JP, Terrin N, Schmid CH, Olkin I. The case of the misleading funnel plot. *BMJ* 2006;333:597–600.
21. Schulz KF, Chalmers I, Hayes RJ, Altman DG. Empirical evidence of bias: dimensions of methodological quality associated with estimates of treatment effects in controlled trials. *JAMA* 1995;273: 408–412.
22. Schunemann HJ, Jaeschke R, Cook DJ, Bria WF, El-Solh AA, Ernst A, Fahy BF, Gould MK, Horan KL, Krishnan JA, *et al*. An official ATS statement: grading the quality of evidence and strength of recommendations in ATS guidelines and recommendations. *Am J Respir Crit Care Med* 2006;174:605–614.
23. Jones PW. Health status measurement in chronic obstructive pulmonary disease. *Thorax* 2001;56:880–887.
24. Schunemann HJ, Guyatt GH, Commentary–goodbye M. (C)ID! Hello MID, where do you come from? *Health Serv Res* 2005;40:593–597.
25. Schunemann HJ, Puhan M, Goldstein R, Jaeschke R, Guyatt GH. Measurement properties and interpretability of the Chronic respiratory disease questionnaire (CRQ). *COPD* 2005;2:81–89.
26. Puhan MA, Mador MJ, Held U, Goldstein R, Guyatt GH, Schunemann HJ. Interpretation of treatment changes in 6-minute walk distance in patients with COPD. *Eur Respir J* 2008;32:637–643.
27. Redelmeier DA, Bayoumi AM, Goldstein RS, Guyatt GH. Interpreting small differences in functional status: the Six Minute Walk test in

chronic lung disease patients. *Am J Respir Crit Care Med* 1997;155: 1278–1282.
28. Wilt TJ, Niewoehner D, MacDonald R, Kane RL. Management of stable chronic obstructive pulmonary disease: a systematic review for a clinical practice guideline. *Ann Intern Med* 2007;147:639–653.
29. Calverley PMA, Anderson JA, Celli B, Ferguson GT, Jenkins C, Jones PW, Yates JC, Vestbo J. Salmeterol and fluticasone propionate and survival in chronic obstructive pulmonary disease. *N Engl J Med* 2007;356:775–789.
30. Papanikolaou PN, Christidi GD, Ioannidis JPA. Comparison of evidence on harms of medical interventions in randomized and nonrandomized studies. *CMAJ* 2006;174:635–641.
31. Pitrou I, Boutron I, Ahmad N, Ravaud P. Reporting of safety results in published reports of randomized controlled trials. *Arch Intern Med* 2009;169:1756–1761.
32. Santaguida P, Raina P, Ismaila A. McMaster Quality Assessment Scale of Harms (McHarm) for primary studies [accessed November 2, 2012]. Available from: http://hiru.mcmaster.ca/epc/mcharm.pdf.

# Incorporating Considerations of Cost-Effectiveness, Affordability, and Resource Implications in Guideline Development

## Article 6 in Integrating and Coordinating Efforts in COPD Guideline Development. An Official ATS/ERS Workshop Report

**Suzanne R. Hill, Leslie G. Olson, Yngve Falck-Ytter, Alvaro A. Cruz, David Atkins, Michael Baumann, Roman Jaeschke, Thomas Woitalla, and Holger J. Schünemann; on behalf of the ATS/ERS Ad Hoc Committee on Integrating and Coordinating Efforts in COPD Guideline Development**

***Introduction*: Professional societies, like many other organizations around the world, have recognized the need to use rigorous processes to ensure that health care recommendations are based on the best available research evidence. This is the sixth of a series of 14 articles prepared to advise guideline developers for respiratory and other diseases on how to achieve this goal. In this article, we focused on integrating cost and resource information in guideline development and formulating recommendations focusing on four key questions.**

***Methods*: We addressed the following specific questions. (*1*) When is it important to incorporate costs, and/or resource implications, and/or cost-effectiveness, and/or affordability considerations in guidelines? (*2*) Which costs and which resource use should be considered in guidelines? (*3*)What sources of evidence should be used to estimate costs, resource use, and cost-effectiveness? (*4*) How can cost-effectiveness, resource implications, and affordability be taken into account explicitly? Our work was based on a prior review on this topic and our conclusions are based on available evidence, consideration of what guideline developers are doing, and workshop discussions.**

***Results and Discussion*: Many authorities suggest that there is a need to include explicit consideration of costs, resource use, and affordability during guideline development. Where drug use is at issue, "explicit consideration" may need to involve only noting whether the price (easily determined and usually the main component of "acquisition cost") of a drug is high or low. Complex interventions such as rehabilitation services are to a greater degree setting- and system-dependent. Resources used, and the costs of those resources, will vary among systems, and formal identification by a guideline group of the resource requirements of a complex intervention is essential. A clinical guideline usually contains multiple recommendations, and in some cases there are hundreds. Defining costs and resource use for all of them—especially for multiple settings—is unlikely to be feasible. At present, disaggregated resource utilization accompanied by some cost information seems to be the most promising approach. The method for assigning values to costs, including external or indirect cost (such as time off work), can have a significant impact on the outcome of any economic evaluation. The perspective that the guideline assumes should be made explicit. Standards for evidence for clinical data are usually good-quality trials reporting a relevant endpoint that should be summarized in a systematic review. Like others, we are therefore proposing that the ideal sources of evidence for cost and resource utilization data for guideline development are systematic reviews of randomized controlled trials that report resource utilization, with direct comparisons between the interventions of interest.**

## INTRODUCTION

Professional societies, like many other organizations around the world, have recognized the need to use more rigorous processes to ensure that health care recommendations are based on the best available research evidence. Most "guidelines for guidelines" documents recommend that consideration of costs and resource allocation should be included in the process of guideline development. Most do not, however, give a detailed account of the justification for doing so, or exactly what is meant by "consideration of costs" being "included in the process."

An example of the uncertainties that may result is the recommendation in chronic obstructive pulmonary disease (COPD) guidelines that "long-acting bronchodilators" should be used in the treatment of moderate and severe COPD (1). Tiotropium is one of the long-acting bronchodilators; treating one patient costs, roughly, $100 a month (2). Applying estimated COPD (3) prevalence figures and distribution of severity for São Paulo (4) to the population of Brazil gives an affected population of 3.78 million with moderate to severe COPD, so the total cost of using tiotropium would be up to $4.5 billion (or approximately €2.7 billion) per year. If the price of the drug were the same compared with European countries, this would be well over 10% of Brazil's total public health expenditure ($34 billion in the year 2000) and probably unaffordable for universal reimbursement, unless there were good evidence that it would markedly reduce other costs borne by the public health system (5, 6). This coarse cost estimate discloses an impasse: one of the options in the guideline's recommendation is, at first glance, probably unaffordable in a middle-income country such as Brazil, but because the guideline's endorsement of tiotropium does not include cost considerations, a more nuanced assessment of the recommendation's viability for low- to middle-income countries is challenging. One could, therefore, assume that a disaggregated presentation of available data on costs and resource use for various treatment options, not just a global assessment of "cost-effectiveness," is needed to inform the development of recommendations.

In June 2007 the American Thoracic Society (ATS) and the European Respiratory Society (ERS) convened an international workshop of methodologists and researchers from around the world for coordinating efforts in guideline development for

---

This article is a section of "Integrating and Coordinating Efforts in Chronic Obstructive Pulmonary Disease (COPD) Guideline Development," an American Thoracic Society (ATS) and European Respiratory Society (ERS) Workshop Report. This official ATS/ERS Workshop Report was adopted by the ATS Board of Directors, August 2012, and by the ERS Executive Committee, February 2012.

*Disclaimer*: S.R.H. is a former staff member of the World Health Organization and is now Chair of the Australian Pharmaceutical Benefits Advisory Committee. The authors alone are responsible for the views expressed in this publication and they do not necessarily represent the decisions, policy, or views of the World Health Organization.

Funding for this conference was made possible (in part) by 1R13HL 90485–01 from the National Institutes of Health. The views expressed in written conference materials or publications and by speakers and moderators do not necessarily reflect the official policies of the Department of Health and Human Services; nor does mention of trade names, commercial practices, or organizations imply endorsement by the U.S. Government.

**Proc Am Thorac Soc Vol 9, Iss. 5, pp 251–255, Dec 15, 2012**

**DOI: 10.1513/pats.201208-059ST**
**Internet address: www.atsjournals.org**

COPD and other respiratory diseases (7). This is the sixth of a series of 14 articles prepared to advise guideline developers in respiratory and other diseases. In this article we focus on integrating cost and resource information in the context of respiratory disease guidelines.

## METHODS

In this article we addressed the questions in Table 1. Key questions were developed from previous reviews of methods for incorporating economic considerations in clinical guidelines and through discussion among the authors (8–10). The review by Edejer (8) was used as the main source for methodological literature, together with the report of the evaluation of the clinical guidelines program of the National Institute for Clinical and Public Health Excellence (11). For illustrative purposes, some published cost-effectiveness analyses of interventions recommended in respiratory guidelines were identified, but a systematic review of such studies was not undertaken.

## RESULTS

### 1. When Is It Important to Incorporate Costs, and/or Resource Implications, and/or Cost Effectiveness, and/or Affordability Considerations in Guidelines?

Most recent recommendations about methods for guideline development say that there is a need to include explicit consideration of costs, resource use, and affordability. In some cases, however, "explicit consideration" may need to involve only noting whether the price (easily determined and usually the main component of "acquisition cost") of a drug is high or low. In other cases, a formal cost-effectiveness analysis may be a useful component in the development of a guideline. In general, the more narrowly focused a guideline is, and the simpler the interventions being considered (e.g., a single pharmacological intervention), the less the need to consider cost and resource use in any detail. Examples of narrowly focused guidelines developed and published by the World Health Organization (WHO) are recommendations on avian influenza or postpartum hemorrhage (12, 13). In these cases, where the core decision follows from the evaluation of one drug or drug combination versus another, the relative acquisition costs of the pharmaceuticals are the important economic consideration.

Formal cost-effectiveness analyses in these circumstances may be needed in some guidelines, but they are only occasionally likely to be helpful. There are several reasons for this.

The first reason is that formal estimates of the cost-effectiveness of a pharmaceutical are generally dependent on the acquisition price and its effectiveness relative to alternatives. If a product is much more expensive than alternatives but offers little added benefit in high-quality clinical trials, a valid formal cost-effectiveness analysis is likely to confirm what can be worked out using a "back of an envelope" analysis based on number needed to treat (NNTs): that the drug is expensive in relation to what it does. Formal cost-effectiveness analyses that suggest different conclusions should be approached with caution.

**TABLE 1. QUESTIONS ADDRESSED REGARDING COST AND RESOURCE IMPLICATIONS IN GUIDELINE DEVELOPMENT**

1. When is it important to incorporate costs, and/or resource implications, and/or cost-effectiveness, and/or affordability considerations in guidelines?
2. Which costs and which resource use should be considered in guidelines?
3. What sources of evidence should be used to estimate costs, resource use, and cost-effectiveness?
4. How can cost-effectiveness, resource implications, and affordability be taken into account explicitly?

For guideline panels, this information could lead to explicit labeling of recommendations that are particularly sensitive to resource considerations.

The second reason is that the methods for calculating community and health care sector costs can be uncertain. Values assigned to these costs often have an element of arbitrariness and are susceptible to bias despite many efforts to establish and ensure appropriate methodological rigor (13). Even if the assigned costs are valid in the system studied, their applicability to other settings is limited. Disaggregated resource use is more likely to be informative.

Third, assessing the validity of a cost-effectiveness analysis is complex, time consuming, and often impossible from the published reports (*see below*). This need not deter guideline panels from making recommendations based on cost-effectiveness analyses, but it should lead to explicit labeling of recommendations that are particularly sensitive to resource considerations.

A well-done formal economic evaluation might inform decision-making, but may also be difficult to interpret unless the decision makers have an established system for bench-marking estimates of cost-effectiveness ratios to give them meaning. That is, a formal economic evaluation might validly conclude that a new drug used for one year prevents x adverse outcomes at a cost of y dollars, but whether y/x counts as "cost-effective" is a value judgment. At present, only three national systems have published indicative thresholds for acceptable cost effectiveness based on incremental cost per quality-adjusted life year (or similar metric), and generalizing this experience to other settings is difficult (14, 15).

Complex interventions such as rehabilitation services are even more setting- and system-dependent than pharmaceuticals. Resource use due to complex interventions will vary among systems. Formal identification of the resources used and consideration of the results by a guideline group is essential. Accurate and complete identification of costs is also important because these data bear on whether the intervention should be recommended for all settings. Disaggregated resource use needs to be reported because identifying resource consequent to a complex intervention is the key to assessing the applicability of the results from one setting to others, and to assessing the practicability of the intervention in different settings. For example, if a rehabilitation program requires two full-time physical therapists, but an institution has four vacant positions for physical therapists that it cannot fill although money is available, the "cost-effectiveness" of the program is irrelevant.

An additional reason for why cost-effectiveness analyses are of limited usefulness is that the decisions that system managers make about the affordability of an intervention are generally influenced more by its total cost rather than by the any estimate of cost-effectiveness ratio (as in the example from Brazil with which we began). As a result, interventions for common diseases may be less affordable for the system than equally cost-effective interventions for uncommon diseases. In the United Kingdom, for example, national spending on health has increased (as a proportion of gross domestic product) over the past 10 years. Guideline development groups in the United Kingdom have recommended interventions on the basis of favorable cost-effectiveness (for example, nasal CPAP for moderate and severe obstructive sleep apnea) but, at the regional level, where there has been an effective cut in funds, these national recommendations have not been implemented because the total costs are locally unaffordable when a disease is common.

### 2. Which Costs and Which Resource Use Should Be Considered in Clinical Guidelines?

A guideline usually contains multiple recommendations. Defining costs and resource use for all of them, especially for multiple settings, is unlikely to be feasible. Recent experience in WHO

guidelines has highlighted the difficulty of presenting costs and resource use in a useful way. At present, disaggregated resource utilization, accompanied by some cost information, seems to be the most promising. However, an important issue is what costs to include. Costs have been considered as "direct" and "indirect" costs, but a more useful strategy may be to specify the nature of costs incurred (i.e., patient costs), community costs, and societal costs (16). The method for assigning currency values to costs (such as time off work, and whether these are classified as patient costs or societal costs) can have a significant impact on the outcome of any economic evaluation. For example, a published economic evaluation of tiotropium for the treatment of COPD claims to show that there are economic benefits associated with using tiotropium compared with ipratropium (17). This conclusion depends, in part, on including both hospital costs and the costs of "inactivity days," as well as the precise dollar values assigned to those costs. In settings other than the one in which the study was performed, or with a different perspective on the setting in which it was performed, one or both of the costs might be considered of less importance. The dollar values assigned to them might be quite different, or they might be incurred in different parts of the system.

Which costs are included is related to the point of view (perspective) of the evaluation, which should be defined. For guideline panels, the relevant perspective is normally assumed to be that of the health system (national or local) in which the guideline would be implemented, but this should be made explicit. The guideline panel should also state whether it has considered the possibility that an intervention that reduces costs at one level in a health system may increase costs to another level of the health system or in a sector outside the health system.

In formulating recommendations, the choice of the comparator intervention is crucial for evaluating data on costs. In the case of tiotropium, the comparator chosen for the formal economic evaluation was ipratropium, based on one trial (18), although placebo-controlled trials were available. This choice reduces the calculated incremental cost of tiotropium. Table 2 shows the relevant data from two of these trials, which suggest that there were significant differences in the control event rate, so that the difference in the percentage of patients with exacerbations was well over 10% between tiotropium and ipratropium, and less than 5% between tiotropium and placebo (18, 19). An economic evaluation based on all trials may well have resulted in different estimates of cost-effectiveness.

### 3. What Sources of Evidence Should Be Used by Guidelines Groups to Estimate Costs, Resource Use, and Cost-Effectiveness?

Standards for evidence for clinical data are usually a systematic review of good-quality trials reporting a relevant clinical endpoint. Standards for economic evaluation have also been described, but are generally less well adhered to (20, 21). To date, several guideline methods manuals have suggested that systematic reviews of cost-effectiveness analyses should also be conducted as part of the process of guidelines development, but the value of this approach needs to be considered carefully.

**TABLE 2. SUMMARY OF CLINICAL TRIAL DATA (FROM REFERENCES 18 AND 19)**

| | Placebo-controlled Trial | | Ipratropium-controlled Trial | |
|---|---|---|---|---|
| | Placebo | Tiotropium | Ipratropium | Tiotropium |
| Percentage of subjects with exacerbation(s) | 32.3 | 28.0 | 53.5 | 39.9 |

First, published cost-effectiveness analyses based on decision-analytic models or related techniques often do not contain enough information to allow their validity to be assessed. In part this reflects the limitations of publishing manuscripts with limits of 2,000 to 3,000 words without full description of supporting electronic models. However, it is also a reflection of the fact that economic models are a synthesis of data from many sources, requiring judgments about selection of variables included and the values assigned to them. These judgments are subject to bias and error. Without access to the model used, it is not possible to assess the validity of the results or conclusions. Evaluation of electronic models is time-consuming and requires highly developed skills, so that a thorough evaluation of cost-effectiveness analyses may be beyond most guideline groups.

Second, published analyses are generally applicable only where and when they were performed, and their lack of transferability has been well documented (22). Undertaking a systematic review of a large number of cost-effectiveness analyses done in different places at different times is not likely to be informative. The main value of systematic review of cost-effectiveness evaluations would be the identification of the key variables that seem likely to determine the cost and affordability of an intervention in the specific context to which the guideline group's work is relevant.

An alternative standard of evidence might be economic evaluations performed alongside clinical trials. There is an increasing number of published studies of this type (23, 24). Where the trial is a large multicenter trial, particularly if it involves countries from different income levels, there is potential to generalize accurately. However, studies of the latter type are still unusual and selective publication of the economic component of trials tends to be the exception. For complex interventions, there may be multiple randomized trials with economic evaluations, each from a single setting, and it is possible that a systematic review of these studies would be informative if the trial settings were sufficiently homogeneous, but this is a question for future research (25).

We are therefore proposing that the ideal sources of evidence for cost and resource utilization data for consideration by guidelines groups are systematic reviews of randomized controlled trials that report resource utilization, with direct comparisons between the interventions of interest. The "interventions of interest" should be chosen from a defined and carefully considered point of view, and should not simply be those for which data happen to be available.

Direct or head-to-head comparison is essential, and indirect comparisons of interventions using separate placebo-controlled trials are not a substitute (25). The reason is that the assessment of comparative resource use across different trials is confounded by too many uncertainties to make such evaluations reliable. We suggest that if there are no head-to-head randomized controlled trials measuring resource use, direct nonrandomized comparisons of the interventions in the same setting (i.e., before-and-after studies) may be a better alternative source of evidence than indirect comparisons. The issue is the consistency of factors such as patient mix, workloads, morale, the expertise of staff, management policies, and so on, and we guess that these are probably more consistent within an institution over (a relatively short) time than between institutions. Whether approaches such as network meta-analyses are of use in cost analyses remains to be explored, but they could provide additional, albeit indirect, data.

### 4. How Can Cost-Effectiveness, Resource Implications, and Affordability Be Taken into Account Explicitly?

The mechanics of considering cost-effectiveness, resource implications, and affordability in making recommendations are challenging. The experience of the National Institute for Health and

**TABLE 3. RECOMMENDATIONS FOR FURTHER RESEARCH**

- What is the relationship between the incorporation of costs into a guideline and the cost impact of a guideline? What are the optimum methods for using cost data in guideline development and of assessing the cost impact of a guideline? Should these processes be unified or separate?
- What are the implications for level of evidence and strength of recommendation taxonomies of considering a range of treatment attributes beyond effectiveness and tolerability?
- What is the role of decision analysis in the development of clinical guidelines?
- In what circumstances is it necessary to use formal consensus methods within a guideline development process?

Reprinted by permission from Reference 25.

Clinical Excellence (NICE) is one of the most systematic attempts, and it is clear that there is significant variability in the way NICE guideline development groups consider these factors (11), even for the single national system for which NICE provides advice. At a global level, the problem is more challenging still.

The mechanism we propose emphasizes the detailed evaluation of trial data with minimal economic modeling. Guideline panels should use disaggregated estimates of resource use and acquisition costs, and weigh these directly against trial evidence of the benefits and harms of the intervention. Disaggregating data on resource use allows each health system to understand the local economic implications of an intervention. For a health system, implementing a guideline may increase spending on pharmaceuticals, but reduce spending on in-patient stays and procedures. However, reducing in-patient stays due to one condition may not save money if demand related to other conditions is very high. Disaggregated data can allow users to consider the effects of different prices for specific resources in their health system.

When guidelines are directed at a global audience there needs to be capacity to adapt recommendations to local conditions. If disaggregated cost data and trial estimates of benefit and harm are included in guidelines, countries with varying levels of income can adapt the recommendations to their circumstances, as can subnational or local groups. It will be critical to define to which setting and under which circumstances a particular recommendation applies.

This does not mean that large amounts of data need necessarily to be assessed or included in guidelines. We noted above the difficulties raised by a hypothetical Brazilian assessment of the GOLD recommendation to use tiotropium for moderate and severe COPD, but the data on frequency of exacerbations shown in Table 2, combined with the cost of the medication, appear sufficient to reach a reliable conclusion. The key is the provision of disaggregated data and the inclusion of all available trials, because it is that which allows each country or local group to substitute their own estimates of acquisition costs and also to make judgments about the transferability of the intervention and the evidence.

## CONCLUSIONS

Although recommended in most guidelines for guidelines to date, there is a significant gap in knowledge to be completed before we can be confident about how best to consider costs in guideline development. This was laid out in 2001 by Eccles and Mason (25). At present, we need to ensure disaggregated presentation of resource utilization and costs and an audit trail for decisions. As described by Eddy, "health interventions are not free, people are not infinitely rich, and the budgets are limited. For every dollar's worth of health care that is consumed, a dollar will be paid. While these payments can be laundered, disguised, or hidden, they will not go away" (26).

***Author Disclosures***: S.R.H. was employed by the World Health Organization. L.G.O. and Y.F.-Y. reported no commercial interests relevant to subject matter. A.A.C. was employed by the World Health Organization. D.A. and M.B. reported no commercial interests relevant to subject matter. R.J. was a member of the GRADE Working Group. T.W. reported no commercial interests relevant to subject matter. H.J.S. was a member of the GRADE Working Group, which received honoraria and lecture fees regarding GRADE. All workshop participants disclosed receipt of honoraria and travel reimbursement from the American Thoracic Society.

### References

1. Rabe KF, Hurd S, Anzueto A, Barnes PJ, Buist SA, Calverley P, Fukuchi Y, Jenkins C, Rodriguez-Roisin R, van Weel C, *et al*. Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease: GOLD executive summary. *Am J Respir Crit Care Med* 2007;176:532–555.
2. Rutten-van Molken MP, Oostenbrink JB, Miravitlles M, Monz BU. Modelling the 5-year cost effectiveness of tiotropium, salmeterol and ipratropium for the treatment of chronic obstructive pulmonary disease in Spain. *Eur J Health Econ* 2007;8:123–135.
3. Halbert RJ, Natoli JL, Gano A, Badamgarav E, Buist AS, Mannino DM. Global burden of COPD: systematic review and meta-analysis. *Eur Respir J* 2006;28:523–532.
4. Menezes AM, Perez-Padilla R, Jardim JR, Muino A, Lopez MV, Valdivia G, Montes de Oca M, Talamo C, Hallal PC, Victora CG. Chronic obstructive pulmonary disease in five Latin American cities (the PLATINO study): a prevalence study. *Lancet* 2005;366:1875–1881.
5. Lima CR, Carvalho MS, Schramm JM. *Cad Saude Publica* 2006;22:1855–1864. (government funding for health and the reliability of national databases in Brazil, 2001–2002).
6. NCEP. Executive summary of the third report of the National Cholesterol Education Program (NCEP) expert panel on detection, evaluation, and treatment of high blood cholesterol in adults (adult treatment panel III). *JAMA* 2001;285:2486–2497.
7. Schunemann HJ, Woodhead M, Anzueto A, Buist S, Macnee W, Rabe KF, Heffner J. A vision statement on guideline development for respiratory disease: the example of COPD. *Lancet* 2009;373:774–779.
8. Edejer TT. Improving the use of research evidence in guideline development: 11. Incorporating considerations of cost-effectiveness, affordability and resource implications. *Health Res Policy Syst* 2006;4:23.
9. Saha S, Hoerger TJ, Pignone MP, Teutsch SM, Helfand M, Mandelblatt JS. The art and science of incorporating cost effectiveness into evidence-based recommendations for clinical preventive services. *Am J Prev Med* 2001;20:36–43.
10. Guyatt G, Baumann M, Pauker S, Halperin J, Maurer J, Owens DK, Tosteson AN, Carlin B, Gutterman D, Prins M, *et al*. Addressing resource allocation issues in recommendations from clinical practice guideline panels: suggestions from an American College of Chest Physicians task force. *Chest* 2006;129:182–187.
11. De Joncheere K, Hill S, Klazinga N, Makela M, Oxman AD. The clinical guideline programme of the National Institute for Public Health and Clinical Excellence. Geneva: World Health Organization; 2007.
12. Schunemann HJ, Hill SR, Kakad M, Vist GE, Bellamy R, Stockman L, Wisloff TF, Del Mar C, Hayden F, Uyeki TM, *et al.* Transparent development of the WHO rapid advice guidelines. *PLoS Med* 2007;4:e119.
13. World Health Organization. WHO recommendations for the prevention of postpartum haemorrhage. Geneva: WHO; 2007.
14. Henry DA, Hill SR, Harris A. Drug prices and value for money: the Australian Pharmaceutical Benefits Scheme. *JAMA* 2005;294:2630–2632.
15. Diamond GA, Forrester JS, Hirsch M, Staniloff HM, Vas R, Berman DS, Swan HJ. Application of conditional probability analysis to the clinical diagnosis of coronary artery disease. *J Clin Invest* 1980;65:1210–1221.
16. Drummond M, Sculpher MJ, Torrance GW, O'Brien BJ, Stoddart GL. Methods for the economic evaluation of health care programmes, 3rd ed. Oxford, UK: Oxford University Press; 2005.
17. Oostenbrink JB, Rutten-van Molken MP, Al MJ, Van Noord JA, Vincken W. One-year cost-effectiveness of tiotropium versus ipratropium to treat chronic obstructive pulmonary disease. *Eur Respir J* 2004;23:241–249.
18. Niewoehner DE, Rice K, Cote C, Paulson D, Cooper JA Jr, Korducki L, Cassino C, Kesten S. Prevention of exacerbations of chronic obstructive pulmonary disease with tiotropium, a once-daily inhaled

anticholinergic bronchodilator: a randomized trial. *Ann Intern Med* 2005;143:317–326.
19. Drummond M, Brown R, Fendrick AM, Fullerton P, Neumann P, Taylor R, Barbieri M. Use of pharmacoeconomics information: report of the ISPOR task force on use of pharmacoeconomic/health economic information in health-care decision making. *Value Health* 2003;6:407–416.
20. Drummond M, Manca A, Sculpher M. Increasing the generalizability of economic evaluations: recommendations for the design, analysis, and reporting of studies. *Int J Technol Assess Health Care* 2005;21:165–171.
21. Sculpher MJ, Drummond MF. Analysis sans frontieres: can we ever make economic evaluations generalisable across jurisdictions? *Pharmacoeconomics* 2006;24:1087–1099.
22. Marson AG, Al-Kharusi AM, Alwaidh M, Appleton R, Baker GA, Chadwick DW, Cramp C, Cockerell OC, Cooper PN, Doughty J, *et al*. The SANAD study of effectiveness of valproate, lamotrigine, or topiramate for generalised and unclassifiable epilepsy: an unblinded randomised controlled trial. *Lancet* 2007;369:1016–1026.
23. Simon J, Gray A, Duley L. Cost-effectiveness of prophylactic magnesium sulphate for 9996 women with pre-eclampsia from 33 countries: economic evaluation of the magpie trial. *BJOG* 2006;113: 144–151.
24. Goldstein RS, Gort EH, Guyatt GH, Feeny D. Economic analysis of respiratory rehabilitation. *Chest* 1997;112:370–379.
25. Eccles M, Mason J. How to develop cost-conscious guidelines. *Health Technol Assess* 2001;5:1–69.
26. Eddy D. A manual for assessing health practices and designing practice policies: The explicit approach. Philadelphia: American College of Physicians; 1992.

# Synthesis, Grading, and Presentation of Evidence in Guidelines

## Article 7 in Integrating and Coordinating Efforts in COPD Guideline Development. An Official ATS/ERS Workshop Report

**Gordon Guyatt, Elie A. Akl, Andy Oxman, Kevin Wilson, Milo A. Puhan, Timothy Wilt, David Gutterman, Mark Woodhead, Elliott M. Antman, and Holger J. Schünemann; on behalf of the ATS/ERS Ad Hoc Committee on Integrating and Coordinating Efforts in COPD Guideline Development**

***Introduction*: Professional societies, like many other organizations around the world, have recognized the need to use more rigorous processes to ensure that health care recommendations are informed by the best available research evidence. This is the seventh of a series of 14 articles that were prepared to advise guideline developers in respiratory and other diseases on approaches for guideline development. This article focuses on synthesizing, rating, and presenting evidence in guidelines.**

***Methods*: In this review we addressed the following questions. (*1*) What evidence should guideline panels use to inform their recommendations? (*2*) How should they rate the quality of the evidence they use? (*3*) How should they grade evidence regarding diagnostic tests? (*4*) What should they do when quality of evidence differs across outcomes? (*5*) How should they present the evidence in a guideline? We did not conduct systematic reviews ourselves. We relied on prior evaluations of electronic databases and systematic reviews suggesting that the Grades of Recommendation, Assessment, Development and Evaluation Working Group (GRADE) approach includes the desired features of a system for grading quality of evidence, including provision of models for presenting evidence for guideline panels, and for the consumers of practice guidelines. This article describes the GRADE approach to grading the quality of evidence and presenting evidence. Available evidence, the practice of leading guideline developers, and workshop discussions provide the basis for our conclusions.**

***Results and Discussion*: GRADE rates the quality of evidence for each outcome across studies rather than for each study. In the GRADE approach randomized trials start as high-quality evidence and observational studies as low-quality evidence, but both can be rated down or up. Five factors may lead to rating down the quality of evidence: study limitations or risk of bias, inconsistency of results, indirectness of evidence, imprecision, and publication bias. Three factors may lead to rating up the quality of evidence from observational studies: large magnitude of effect, dose–response gradient, and situations in which all plausible confounders would decrease an apparent treatment effect, or would create a spurious effect when results suggest no effect. GRADE suggests use of evidence profiles that provide a comprehensive way to display the key evidence relevant to a clinical question. Guideline developers who follow this structure will find the transparency of their recommendations markedly enhanced.**

## INTRODUCTION

Professional societies, like many other organizations around the world, have recognized the need to use more rigorous processes to ensure that health care recommendations are informed by the best available research evidence. In June 2007, the American Thoracic Society (ATS) and the European Respiratory Society (ERS) convened an international workshop of methodologists and researchers from around the world to coordinate efforts in guideline development using chronic obstructive pulmonary disease (COPD) as a model (1). Participants completed the work during the subsequent four years to develop a series of recommendations. This is the seventh of a series of 14 articles that were prepared to advise guideline developers in respiratory and other diseases on approaches for guideline development. This article focuses on synthesis, grading, and presentation of evidence in guidelines.

## METHODS

The authors of this article developed and discussed the key questions in this article, which are listed in Table 1. We did not conduct systematic reviews ourselves. We relied on prior evaluations of electronic databases and systematic reviews suggesting that the Grades of Recommendation, Assessment, Development and Evaluation Working Group (GRADE) approach includes the desired features of a grading system (2). The GRADE working group has also prepared models for presenting evidence for guideline panels (3, 4). Thus, this article focuses on the description of the GRADE approach to grading the quality of evidence and presenting evidence. Our conclusions are based on available evidence, the practice of leading guideline developers, and workshop discussions.

## RESULTS

### 1. What Evidence Should Guideline Panels Use to Inform their Recommendations?

Most authorities accept that clinical guidelines should be evidence-based. Although consensus about what exactly that means may be less complete, all would agree that an attempt to collect and succinctly summarize the evidence regarding the impact of alternative management strategies on all patient-important outcomes is a prerequisite. The widely used term for this collection and summarization is "systematic review." Most high-profile organizations that produce guidelines endeavor to provide their panels with relevant systematic reviews—usually already available in the literature, but sometimes prepared for the guideline itself (5–7).

Systematic reviews involve clear definition of the question (including the patients, intervention, comparator, and outcome), restriction to the highest-quality evidence available (typically randomized trials),

This article is a section of "Integrating and Coordinating Efforts in Chronic Obstructive Pulmonary Disease (COPD) Guideline Development," an American Thoracic Society (ATS) and European Respiratory Society (ERS) Workshop Report. This official ATS/ERS Workshop Report was adopted by the ATS Board of Directors, August 2012, and by the ERS Executive Committee, February 2012.

Funding for this conference was made possible (in part) by 1R13HL 90485-01 from the National Institutes of Health. The views expressed in written conference materials or publications and by speakers and moderators do not necessarily reflect the official policies of the Department of Health and Human Services; nor does mention of trade names, commercial practices, or organizations imply endorsement by the U.S. Government.

Holger Schünemann holds the Michael Gent Chair in Healthcare Research.

**Proc Am Thorac Soc Vol 9, Iss. 5, pp 256–261, Dec 15, 2012**

**DOI: 10.1513/pats.201208-060ST**
**Internet address: www.atsjournals.org**

**TABLE 1. QUESTIONS ADDRESSED REGARDING SYNTHESIS, GRADING, AND PRESENTATION OF EVIDENCE IN GUIDELINES**

| |
|---|
| 1. What evidence should guideline panels use to inform their recommendations? |
| 2. How should they rate the quality of the evidence they use? |
| 3. How should they grade evidence regarding diagnostic tests? |
| 4. What should they do when quality of evidence differs across outcomes? |
| 5. How should they present the evidence in a guideline? |

a comprehensive search for eligible studies, and an unbiased summary of the evidence from those trials (often a statistical summary presented in a meta-analysis). In the COPD area the patient population would typically be a heterogenous group of patients with COPD, and patient-important outcomes would include mortality, exacerbations, hospitalization, and health-related quality of life. Use of a high-quality systematic review ensures that the definition of the question, eligibility criteria, the search strategy, and the evaluation of study quality and results are all transparent.

Guideline panels should be using systematic reviews of high quality. Individual articles and whole texts describing the process of conducting a systematic reviews are available, including guidance from the Cochrane Collaboration (8). Should they wish to assess the quality of available systematic reviews, guideline panels can use one of a number of available instruments (9–12).

Panelists may have more than one systematic review available to them. Guidance is available for locating such reviews, judging their eligibility, incorporating evidence from the reviews, and reporting the results of the process (13). Particularly if they have a statistician available to conduct an updated meta-analysis, panels can also update prior meta-analyses with the most recent studies.

Unfortunately, some questions related to COPD management have not been addressed in randomized trials. As a result, the highest-quality evidence may be of low or even very low quality. Use of the GRADE approach is completely compatible with use of even very low–quality evidence.

## 2. How Should Guideline Panels Rate the Quality of the Evidence they Use?

While guideline panels require complete and detailed evaluations of the underlying evidence, guideline users need succinct summaries of evidence that include ratings of quality. Clinicians and patients need to know whether estimates of benefit and risk are based on high-quality evidence—and are therefore unlikely to change as further evidence accumulates—or find support from only low-quality evidence, and are therefore insecure and likely to change.

Responding to this need, most organizations developing guidelines use structured approaches to rating quality of evidence (7, 14–19). The Canadian Optimal Medication Prescribing and Utilization Service (COMPUS), a nationally coordinated program funded by Health Canada and delivered by the Canadian Coordinating Office for Health Technology Assessment, has conducted the most recent and comprehensive review of existing systems (9). COMPUS assembled a working group of internal researchers, information specialists, methodology experts, and external researchers who evaluated over 50 evidence grading systems. Using evaluation goals developed by the U.S. Agency for Healthcare Research and Quality (10) the experts gave the highest rating to the GRADE (20) and the Scottish Intercollegiate Guideline Network (14) systems. A second round of expert consultation and stakeholder input from all interested parties confirmed the selection of these instruments. The Scottish Intercollegiate Guideline Network has subsequently adopted the GRADE approach.

GRADE is unique in the depth and breadth of methodological guideline development and clinical expertise devoted to its development, its explicitness in criteria for rating quality of evidence, and its wide acceptance. Over 60 organizations involved with summarizing evidence and developing recommendations, including the World Health Organization, UpToDate, Clinical Evidence and the Cochrane Collaboration, have adopted GRADE. Among these organizations are three that are prominent within the respiratory community: the American Thoracic Society (21), the American College of Chest Physicians (22), and the European Society of Thoracic Surgeons. The care with which the system was developed, as well as its rapid and extensive dissemination, recommends its use in COPD guidelines. We will now describe the GRADE approach to rating quality of evidence.

***GRADE's approach to rating evidence quality.*** GRADE rates the quality of evidence for each outcome across studies rather than rating each study as a single unit. The quality of evidence may differ from one outcome to another within a single study (if, for instance, no patients are lost to follow-up for mortality, but status regarding quality of life is unavailable for a large proportion of patients). In the GRADE approach, randomized controlled trials (RCTs) start as high-quality evidence, observational studies as low-quality evidence (Table 2). Five factors may lead to rating down the quality of evidence, three to rating up (Table 2). Ultimately, the quality of evidence for each outcome falls into one of four categories from high to very low. Some groups who have adopted the key elements of GRADE—notably UpToDate and the ACCP—collapse low- and very low–quality evidence into a single category.

***Factors that decrease the quality of evidence.*** The following limitations may decrease the quality of evidence supporting a recommendation (Table 2).

1. Limitation in the study design or execution (risk of bias). Our confidence in recommendations decreases if studies suffer from major limitations that are likely to result in a biased assessment of the treatment effect. These methodological limitations include lack of concealment of randomization, lack of blinding with subjective outcomes highly susceptible to bias, a large loss to follow-up, or RCTs stopped early for benefit. For instance, most of 13 randomized trials of oral bacterial extracts with putative immunomodulatory properties for reducing COPD exacerbations failed to report procedures to ensure concealment of randomization, failed to report how they dealt with patients lost to follow-up, and failed to specify adherence to the intention-to-treat principle (23). These limitations would probably lead guideline panels to downgrade the quality of evidence from high to moderate quality.
2. Unexplained heterogeneity of results (inconsistent results). When studies yield widely differing estimates of the treatment effect (heterogeneity or variability in results), investigators should look for explanations for that heterogeneity. For instance, drugs may have larger relative effects in sicker populations or when given in larger doses. When heterogeneity exists, but investigators fail to identify a plausible explanation, the quality of evidence decreases. For example, four blinded crossover randomized trials of supplemental oxygen in patients with COPD with exercise hypoxemia have suggested little if any improvement in day-to-day dyspnea (24–27), while a fifth randomized trial suggests substantial benefit (28). Apparent differences in patients, interventions, measurement of outcome, or study design and conduct fail to explain the different results. This unexplained variability would mandate the classification of this otherwise high-quality evidence as moderate quality.

**TABLE 2. GRADE APPROACH TO GRADING THE QUALITY OF EVIDENCE**

| Source of Body of Evidence | Initial Rating of Quality | Factors that May Decrease the Quality | Factors that May Increase the Quality | Final Quality of a Body of Evidence |
|---|---|---|---|---|
| Randomized trials | High | 1. Risk of bias | 1. Large effect | High (⊕⊕⊕⊕) |
| | | 2. Inconsistency | 2. Dose–response | Moderate (⊕⊕⊕○) |
| Observational studies | Low | 3. Indirectness | 3. All plausible residual confounding would reduce the demonstrated effect or would suggest a spurious effect if no effect was observed | Low (⊕⊕○○) |
| | | 4. Imprecision | | Very low (⊕○○○) |
| | | 5. Publication bias | | |

3. Indirectness of evidence. There are two types of indirectness of evidence. The first occurs when clinicians may be interested in which of two or more effective agents to prescribe. For instance, clinicians may wonder if different formulations of inhaled steroids impact differently on patients with COPD. Randomized trials may be available, but they may compare each agent to placebo, rather than directly compare the alternatives under clinical consideration. Such indirect comparisons yield only moderate- or low-quality evidence.
   The second type of indirect comparison occurs when the guideline panel has a specific clinical question but investigators may have undertaken studies in similar, but not identical populations; tested similar but not identical interventions against similar but not identical comparators; and measured outcomes that are related to, but do not represent those in which we are primarily interested. Table 3 presents examples of each of these indirect comparisons (4).
4. Imprecision. When studies include few patients and few events and thus have wide confidence intervals, a guideline panel will rate the quality of the evidence lower than it otherwise would because of resulting uncertainty in the results. For instance, three trials in the previously mentioned systematic review of oral bacterial extracts suggested a 1/3 reduction in exacerbations, but the wide confidence interval precludes confidence in the apparent effect (relative risk [RR], 0.66; 95% confidence interval, 0.41–1.08) and would likely lead guideline panels to further downgrade the quality of evidence for this outcome (from moderate to low quality) (23).
5. Publication bias. The quality of evidence may be reduced if investigators fail to report studies, which are typically those that show no effect. Unfortunately, guideline panels must make guesses about the likelihood of publication bias. A prototypical situation that should elicit suspicion of reporting bias is when published evidence includes a number of small trials, all of which are industry funded (29). A "funnel plot" demonstrating larger effects in smaller trials suggests that small trials with minimal effects remain unpublished. Figure 1 provides an example of a skewed distribution of results of small trials from a systematic review of oxygen in COPD reporting on the results of tests of exercise capacity (30). Tests such as funnel plots are, however, of limited use (31); avoiding reporting bias, if possible, is far preferable than trying to detect it.

A particular body of evidence can suffer from more than one of these limitations, and the greater the limitations, the lower the quality of the evidence. One could imagine a situation in which RCTs were available, but all or virtually all of these limitations would be present, and in serious form—very low quality of evidence would result.

***Factors that increase the quality of evidence.*** While well-done observational studies will generally yield low-quality evidence, there are unusual circumstances in which guideline panels classify such evidence as moderate or even high quality (*see* the last column of Table 2). The most common reason for upgrading the quality of evidence is large or very large treatment effects realized over short periods of time. For example, we are confident that the institution of mechanical ventilation in patients about to die of respiratory failure prolongs life in all such patients, and helps prolong life substantially in some of them. Unfortunately, aside from the institution of mechanical ventilation versus no assisted ventilation in patients whose death from respiratory failure is imminent, such situations in COPD are not common.

An even less common reason for rating up quality of evidence is if all plausible confounders and biases from observational studies unaccounted for in the adjusted analysis (that is, all residual confounders) of a rigorous observational study would result in an underestimate of an apparent treatment effect. If, for instance, only sicker patients receive an experimental intervention or exposure, yet they still fare better, it is likely that the actual intervention or exposure effect is even larger than the data suggest.

For instance, unpublished systematic review addressed the effect of condom use on HIV infection among men who have sex with men. The pooled effect estimate of RR from the five eligible observational studies was 0.34 (0.21–0.54) in favor of condom use compared with no condom use. Two of these studies (32, 33) that examined number of partners in those using condoms and not using condoms found that condom users were more likely to have more partners (but did not adjust for this confounding factor in their analyses). Considering number of partners would, if anything, strengthen the effect estimate in favor of condom use.

**TABLE 3. EVIDENCE IS WEAKER IF COMPARISONS ARE INDIRECT**

| Question of Interest | Source of Indirectness |
|---|---|
| Relative effectiveness of formoterol and salmeterol | Indirect comparison: randomized trials have compared formoterol and salmeterol to placebo, but trials comparing formoterol to salmeterol are unavailable. |
| Respiratory rehabilitation in moderately severe COPD | Differences in population: randomized trials have focused on patients with severe COPD |
| Low-intensity home-based respiratory rehabilitation | Differences in intervention: randomized trials have by and large tested hospital-based or intensive home-based respiratory rehabilitation |
| Effect of interventions on quality of life, exacerbations, and mortality | Differences in outcome: many randomized trials in COPD test the effect of interventions on pulmonary function or exercise capacity |

*Definition of abbreviation*: COPD = chronic obstructive pulmonary disease.

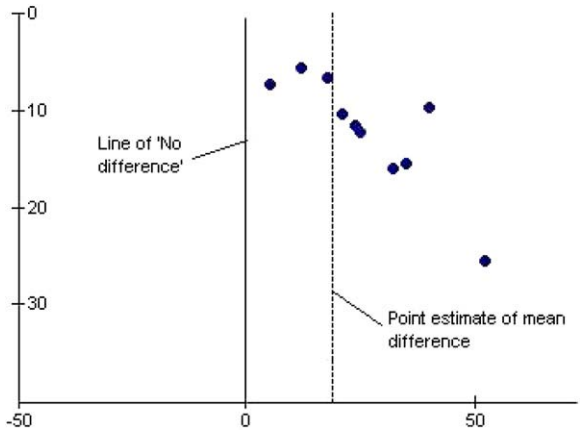


***Figure 1.*** Differences in exercise capacity in (intervention–control groups) in short-term randomized trials of oxygen in patients with chronic obstructive pulmonary disease.

### 3. How Should Guideline Panels Grade Evidence Regarding Diagnostic Tests?

The goal of every management action a clinician takes is to decrease mortality or morbidity, or improve the patient's well-being (34). Consider an 85-year-old woman with end-stage COPD with an apparent lung nodule of unknown duration on plain chest radiograph. A computed tomography (CT) scan shows spiculation consistent with a malignancy but no evident metastases. Other than informing the patient, the CT finding warrants no further action.

The diagnostic test provided additional information, but the patient derived no benefit from the test. Guideline developers must recommend tests on the basis that they improve patient-important outcomes rather than on the basis of their accuracy. Thus, accuracy studies provide evidence for the possible benefit of a test, but that evidence is indirect (34).

Consider patients with COPD presenting to the emergency room with increasing dyspnea and whose clinical presentation suggests the possibility of superimposed heart failure. A number of studies suggest that an elevated B-natriuretic peptide (BNP) markedly increases the likelihood that heart failure is responsible for the patient's symptoms (35–37). What is the quality of evidence supporting a recommendation of use of the test in this setting?

It does not necessarily follow that patients will benefit from application of this accurate test. First, clinicians may be extremely accurate in their assessment without use of the test, which may therefore add nothing to the diagnostic process (34). Second, clinicians may be only moderately accurate, but the test may provide no information beyond what clinicians garner from history, physical examination, and chest radiograph. Finally, the test may improve initial diagnosis, but outcome may not change (for instance, if clinicians ultimately undertake a trial of furosemide that leads to final accurate diagnosis and appropriate treatment). As a result, one could classify evidence regarding test accuracy as indirect, and this indirectness—depending on its extent—could lead to a final quality rating of moderate- or even low-quality evidence.

In many instances, sorting out whether application of a diagnostic test really changes patient outcome requires a controlled trial in which patients are randomized to undergo, or not undergo, the test procedure. In the case of BNP, investigators have conducted a number of such trials (38). One study of 453 patients presenting to an emergency department with dyspnea found that 75% randomized to BNP testing versus 85% randomized to no testing were admitted to hospital ($P = 0.008$) (39). BNP-tested patients who were admitted were discharged from hospital sooner than those not tested (8 vs. 11 d, $P = 0.001$).

This study provides compelling evidence that—with the patient mix tested and in the particular setting—patients tested with BNP spend less time in the hospital, a benefit that most patients would value. A guideline panel would consult not only this single study, but a systematic review of all such studies—which suggests a mortality reduction with use of BNP. The randomized trial design of these studies would suggest a rating of high-quality evidence. The panel may still rate the evidence down to moderate quality for other reasons (they may consider the evidence indirect either because few patients with COPD participated in

**TABLE 4. SUMMARY OF FINDINGS OF ANTICHOLINERGIC AGENTS IN CHRONIC OBSTRUCTIVE PULMONARY DISEASE**

| Quality Assessment | | | | | | | Summary of Findings: No. of Patients | | Summary of Findings: Effect | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| No. of Studies | Design | Limitations | Consistency | Directness | Imprecision | Other Considerations | Anticholinergics | Placebo | Relative (*95% CI*) | Absolute | Quality | Importance |
| Respiratory Deaths (follow-up mean 20 mo) | | | | | | | | | | | | |
| 5 | Randomized trial | No serious limitations* | No serious inconsistency | No serious indirectness | Serious imprecision† | None | 2/4,036 | 12/3,845 2.5%‡ | RR, 0.27 (0.09–0.81) | 1 fewer per 1,000 18 fewer per 1,000[3] | ⊕⊕⊕⊕ HIGH | CRITICAL |
| Exacerbations Resulting in Withdrawal from Study (follow-up 3–6 mo) | | | | | | | | | | | | |
| 6 | Randomized trial | No serious limitations* | No serious inconsistency | No serious indirectness | No serious imprecision† | None | 126/2,516 | 181/2,075 | RR, 0.60 (0.48–0.75) | 30 fewer per 1,000§ | ⊕⊕⊕⊕ HIGH | CRITICAL |
| Hospitalizations due to Exacerbations (follow-up 3–6 mo) | | | | | | | | | | | | |
| 3 | Randomized trial | No serious limitations* | No serious inconsistency | No serious indirectness | No serious imprecision | None | 106/1,866 | 142/1,686 | RR, 0.67 (0.53–0.86) | 20 fewer per 1,000[4] | ⊕⊕⊕⊕ HIGH | CRITICAL |

* Concealment of allocation not clear in several studies.
† There were few events in both groups.
‡ Data from the TORCH trial indicate a 2.5% respiratory death rate over 20 months.
§ Data from the control group of the studies included in the meta-analysis used to calculate absolute effect.
**Author(s):** Nancy Santesso and Holger Schunemann.
**Date:** 2007-06-04.
**Question:** Should anticholinergics vs placebo be used in COPD?
**Settings:** Outpatient.
Data obtained from Referece 42.

the trial or because they question the applicability to other settings), but there is no question that the randomized trials examining the impact of the test on patient-important outcomes strengthen the evidence supporting BNP use in comparison to studies that address only the tests' accuracy.

### 4. What Should Guideline Panels Do When Quality of Evidence Differs across Outcomes?

Recommendations depend on evidence regarding a number of outcomes, and in the approach we have described above, one would establish the quality of evidence for each outcome that the panel deemed important (such as mortality, exacerbations, health-related quality of life, or hospitalizations). Ideally, the decision about importance would include input from patient representatives, patients themselves, or evidence regarding importance from the literature.

One might call the GRADE approach "outcome-centric." That is, GRADE requires assessment of evidence on an outcome-by-outcome basis, and acknowledges that, often, the quality of the evidence will differ across outcomes. For instance, we have some evidence for effectiveness of antibiotics in COPD exacerbations from randomized trials, but evidence regarding rare side effects such as anaphylaxis comes from observational studies. Guideline panels will often find they have high-quality evidence regarding benefits of treatment, but only low-quality evidence regarding harms.

This presents a potential dilemma. How should one rate the overall quality of evidence if quality differs across patient-important outcomes? When, for instance, randomized trials have addressed effectiveness but only observational studies provide evidence regarding toxicity, should the overall quality of evidence be considered high, moderate, or low?

The GRADE approach suggests that guideline developers should consider whether downsides of therapy are critical to the decision regarding the optimal management strategy. If the outcome for which evidence is lower quality is indeed critical for decision making, then the rating of overall quality of the evidence must reflect this lower quality evidence. If the outcome for which evidence is of low or very low quality is important but not critical, the GRADE approach suggests an overall rating reflecting the higher-quality evidence from the critical outcomes.

### 5. How Should Guideline Panels Present the Evidence in a Guideline?

The Conference on Guideline Standardization (COGS) developed an 18-item checklist for the reporting of guidelines (40). The checklist includes the method for synthesizing evidence (how evidence was used to create recommendations, e.g., evidence tables, meta-analysis, decision analysis) and the recommendation grading criteria (the criteria used to rate the quality of evidence that supports the recommendations and the system for describing the strength of the recommendations).

GRADE suggests use of evidence profiles that provide a comprehensive way to display all data relevant to a clinical question (Table 4). Guideline developers and those developing evidence syntheses can use GRADEpro software (http://www.cc-ims.net/gradepro or http://www.flintbox.com/technology.asp?page=3993) to develop these profiles using a comprehensive handbook (41). These evidence profiles include summaries of the information for each critical outcome that influences clinical decision-making, including a detailed evaluation of the study quality by outcome and the associated effects. Such tables have proved invaluable in helping guideline panels develop evidence-based recommendations.

## SUMMARY

Well-conducted systematic reviews are required to produce high-quality clinical practice guidelines. GRADE provides an explicit and comprehensive structure defining the role of study design in determining evidence quality, and delineating five categories of limitations that may lower study quality, and three categories that may raise study quality. Guideline developers who follow this structure will find the transparency of their recommendations markedly enhanced.

***Author Disclosures***: G.G. was a member of the GRADE Working Group. E.A.A. and A.O. reported no commercial interests relevant to subject matter. E.M.A. was chair of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines. K.W. was employed by UpToDate, Inc. M.A.P. received lecture fees from GlaxoSmithKline and Grants from Astra Zeneca, Boehringer Ingelheim and GlaxoSmithKline. T.W. and D.G. reported no commercial interests relevant to subject matter. M.W. consulted with GlaxoSmithKline and received lecture fees from Pfizer. H.J.S. was a member of the GRADE Working Group, which received honoraria and lecture fees regarding GRADE. All workshop participants disclosed receipt of honoraria and travel reimbursement from the American Thoracic Society.

## REFERENCES

1. Schunemann HJ, Woodhead M, Anzueto A, Buist S, Macnee W, Rabe KF, Heffner J. A vision statement on guideline development for respiratory disease: the example of COPD. *Lancet* 2009;373:774–779.
2. Guyatt GH, Oxman AD, Vist GE, Kunz R, Falck-Ytter Y, Alonso-Coello P, Schünemann HJ; GRADE Working Group. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ* 2008;336:924–926.
3. Schünemann H, Oxman AD, Higgins JPT, Vist GE, Glasziou P, Guyatt GH. Chapter 11: Presenting results and "summary of findings"tables. In: GS Higgins JPT, Green S, editors. Cochrane handbook for systematic reviews of interventions, version 5.0.1 [accessed 11 November 2008]. The Cochrane Collaboration; 2008. Available from: www.cochrane-handbook.org.
4. Schünemann HJ, Hill SR, Kakad M, Bellamy R, Uyeki TM, Hayden FG, Yazdanpanah Y, Beigel J, Chotpitayasunondh T, Del Mar C, *et al.*; WHO Rapid Advice Guideline Panel on Avian Influenza. WHO Rapid Advice Guidelines for pharmacological management of sporadic human infection with avian influenza A (H5N1) virus. *Lancet Infect Dis* 2007;7:21–31.
5. Burgers JS, Grol R, Klazinga NS, Makela M, Zaat J; AGREE Collaboration. Towards evidence-based clinical practice: an international survey of 18 clinical guideline programs. *Int J Qual Health Care* 2003;15:31–45.
6. Oxman AD, Schunemann HJ, Fretheim A. Improving the use of research evidence in guideline development: 8. Synthesis and presentation of evidence. *Health Res Policy Syst* 2006;4:20.
7. Harris RP, Helfand M, Woolf SH, Lohr KN, Mulrow CD, Teutsch SM, Atkins D; Methods Work Group, Third US Preventive Services Task Force. Current methods of the US Preventive Services Task Force: a review of the process. *Am J Prev Med* 2001;20:21–35.
8. Higgins J, Green S. Cochrane handbook for systematic reviews of interventions, version 5.0.2 [updated September 2009; accessed 11 November 2008]. The Cochrane Collaboration; 2008. Available from: www.cochrane-handbook.org.
9. Canadian Optimal Medication Prescribing and Utilization Service. COMPUS Evaluation Methodology draft [accessed 11 November 2008]. Available from: https://www.ccohta.ca/compus/compus_pdfs/COMPUS_Evaluation_Methodology_draft_e.pdf.
10. West S, King V, Carey T, Lohr K, McKoy N, Sutton S, Lux L. Systems to rate the strength of scientific evidence. In: Evidence report/technology assessment. Vol. AHRQ Publication No. 02–E016. Rockville, MD: Agency for Healthcare Research and Quality; 2002.
11. Shea BJ, Bouter LM, Peterson J, Boers M, Andersson N, Ortiz Z, Ramsay T, Bai A, Shukla VK, Grimshaw JM. External validation of a measurement tool to assess systematic reviews (AMSTAR). *PLoS ONE* 2007;2:e1350.
12. Shea BJ, Grimshaw JM, Wells GA, Boers M, Andersson N, Hamel C, Porter AC, Tugwell P, Moher D, Bouter LM. Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews. *BMC Med Res Methodol* 2007;7:10.
13. Whitlock EP, Lin JS, Chou R, Shekelle P, Robinson KA. Using existing systematic reviews in complex systematic reviews. *Ann Intern Med* 2008;148:776–782.

14. Harbour R, Miller J. A new system for grading recommendations in evidence based guidelines. *BMJ* 2001;323:334–336.
15. Briss PA, Zaza S, Pappaioanou M, Fielding J, Wright-De Aguero L, Truman BI, Hopkins DP, Mullen PD, Thompson RS, Woolf SH, *et al.* Developing an evidence-based Guide to Community Preventive Services–methods. The Task Force on Community Preventive Services. *Am J Prev Med* 2000;18:35–43.
16. Truman BI, Smith-Akin CK, Hinman AR, Gebbie KM, Brownson R, Novick LF, Lawrence RS, Pappaioanou M, Fielding J, Evans CA Jr, *et al.* Developing the Guide to Community Preventive Services–overview and rationale. The Task Force on Community Preventive Services. *Am J Prev Med* 2000;18:18–26.
17. Albright J, Allman R, Bonfiglio RP, Conill A, Dobkin B, Guccione AA, Hasson S, Russo R, Shekelle P, Susman JL, *et al.* Philadelphia Panel evidence-based clinical practice guidelines on selected rehabilitation interventions: overview and methodology. *Phys Ther* 2001;81:1629–1640.
18. Gibbons RJ, Smith S, Antman E. American College of Cardiology/ American Heart Association clinical practice guidelines: part I. Where do they come from? *Circulation* 2003;107:2979–2986.
19. Ebell MH, Siwek J, Weiss BD, Woolf SH, Susman J, Ewigman B, Bowman M. Strength of recommendation taxonomy (SORT): a patient-centered approach to grading evidence in the medical literature. *J Am Board Fam Pract* 2004;17:59–67.
20. Atkins D, Best D, Briss PA, Eccles M, Falck-Ytter Y, Flottorp S, Guyatt GH, Harbour RT, Haugh MC, Henry D, *et al.*; GRADE Working Group. Grading quality of evidence and strength of recommendations. *BMJ* 2004;328:1490.
21. Schünemann HJ, Jaeschke R, Cook DJ, Bria WF, El-Solh AA, Ernst A, Fahy BF, Gould MK, Horan KL, Krishnan JA, *et al.*; ATS Documents Development and Implementation Committee. An official ATS statement: grading the quality of evidence and strength of recommendations in ATS guidelines and recommendations. *Am J Respir Crit Care Med* 2006;174:605–614.
22. Guyatt G, Gutterman D, Baumann MH, Addrizzo-Harris D, Hylek EM, Phillips B, Raskob G, Lewis SZ, Schünemann H. Grading strength of recommendations and quality of evidence in clinical guidelines: report from an American College of Chest Physicians task force. *Chest* 2006;129:174–181.
23. Steurer-Stey C, Bachmann LM, Steurer J, Tramer MR. Oral purified bacterial extracts in chronic bronchitis and COPD: systematic review. *Chest* 2004;126:1645–1655.
24. Lacasse Y, Lecours R, Pelletier C, Begin R, Maltais F. Randomised trial of ambulatory oxygen in oxygen-dependent COPD. *Eur Respir J* 2005;25: 1032–1038.
25. Liker ES, Karnick A, Lerner L. Portable oxygen in chronic obstructive lung disease with hypoxemia and cor pulmonale. A controlled double-blind crossover study. *Chest* 1975;68:236–241.
26. McDonald CF, Blyth CM, Lazarus MD, Marschner I, Barter CE. Exertional oxygen of limited benefit in patients with chronic obstructive pulmonary disease and mild hypoxemia. *Am J Respir Crit Care Med* 1995;152:1616–1619.
27. Nonoyama M, Brooks D, Guyatt GH, Goldstein RS. Effect of oxygen on health related quality of life in COPD patients with transient exertional hypoxemia. *Am J Respir Crit Care Med* 2007;176:343–349.
28. Eaton T, Garrett JE, Young P, Fergusson W, Kolbe J, Rudkin S, Whyte K. Ambulatory oxygen improves quality of life of COPD patients: a randomised controlled study. *Eur Respir J* 2002;20: 306–312.
29. Bhandari M, Busse JW, Jackowski D, Montori VM, Schunemann H, Sprague S, Mears D, Schemitsch EH, Heels-Ansdell D, Devereaux PJ. Association between industry funding and statistically significant pro-industry findings in medical and surgical randomized trials. *CMAJ* 2004;170:477–480.
30. Bradley J, O'Neill B. Short-term ambulatory oxygen for chronic obstructive pulmonary disease. *Cochrane Database Syst Rev* 2005;4:CD004356.
31. Lau J, Ioannidis JP, Terrin N, Schmid CH, Olkin I. The case of the misleading funnel plot. *BMJ* 2006;333:597–600.
32. Detels R, English P, Visscher BR, Jacobson L, Kingsley LA, Chmiel JS, Dudley JP, Eldred LJ, Ginzburg HM. Seroconversion, sexual activity, and condom use among 2915 HIV seronegative men followed for up to 2 years. *J Acquir Immune Defic Syndr* 1989;2:77–83.
33. Difranceisco W, Ostrow DG, Chmiel JS. Sexual adventurism, high-risk behavior, and human immunodeficiency virus-1 seroconversion among the Chicago MACS-CCS cohort, 1984 to 1992. A case-control study. *Sex Transm Dis* 1996;23:453–460.
34. Schunemann HJ, Oxman AD, Brozek J, Glasziou P, Bossuyt P, Chang S, Muti P, Jaeschke R, Guyatt GH. Grading quality of evidence and strength of recommendations for diagnostic tests and strategies. *BMJ* 2008;336:1106–1110.
35. Hall C. Essential biochemistry and physiology of (NT-pro)BNP. *Eur J Heart Fail* 2004;6:257–260.
36. McCullough PA, Nowak RM, McCord J, Hollander JE, Herrmann HC, Steg PG, Duc P, Westheim A, Omland T, Knudsen CW, Storrow AB, *et al.* B-type natriuretic peptide and clinical judgment in emergency diagnosis of heart failure: analysis from Breathing Not Properly (BNP) Multinational Study. *Circulation* 2002;106:416–422.
37. Maisel AS, Krishnaswamy P, Nowak RM, McCord J, Hollander JE, Duc P, Omland T, Storrow AB, Abraham WT, Wu AH, *et al.* Rapid measurement of B-type natriuretic peptide in the emergency diagnosis of heart failure. *N Engl J Med* 2002;347:161–167.
38. Felker G, Hasselblad V, Hernandez AF, O'Connor CM. Biomarker-guided therapy in chronic heart failure: a meta-analysis of randomized controlled trials. *Am Heart J* 2009;158:422–430.
39. Mueller C, Scholer A, Laule-Kilian K, Martina B, Schindler C, Buser P, Pfisterer M, Perruchoud AP. Use of B-type natriuretic peptide in the evaluation and management of acute dyspnea. *N Engl J Med* 2004;350: 647–654.
40. Shiffman RN, Shekelle P, Overhage JM, Slutsky J, Grimshaw J, Deshpande AM. Standardized reporting of clinical practice guidelines: a proposal from the Conference on Guideline Standardization. *Ann Intern Med* 2003;139:493–498.
41. Schünemann H, Brozek J, Guyatt G, Oxman O, editors. GRADE handbook for grading quality of evidence and strength of recommendation. ed. The GRADE Working Group; 2010.
42. Salpeter SR, Buckley BS, Salpeter EE. Meta-analysis: anticholinergics, but not β-agonists, reduce severe exacerbations and respiratory mortality in COPD. *J Gen Intern Med* 2006;21:1011–1019.

# Integrating Values and Consumer Involvement in Guidelines with the Patient at the Center

## Article 8 in Integrating and Coordinating Efforts in COPD Guideline Development. An Official ATS/ERS Workshop Report

**Marcia Kelson, Elie A. Akl, Hilda Bastian, Françoise Cluzeau, J. Randall Curtis, Gordon Guyatt, Victor M. Montori, Sandy Oliver, and Holger J. Schünemann; on behalf of the ATS/ERS Ad Hoc Committee on Integrating and Coordinating Efforts in COPD Guideline Development**

***Introduction*: Professional societies, like many other organizations around the world, have recognized the need to use rigorous processes to ensure that healthcare recommendations are informed by the best available research evidence. They are also realizing the need to involve consumers of healthcare (patients, caregivers, and the public) and integrate their values and preferences in clinical guideline development. This is the eighth of a series of 14 articles that were prepared to advise guideline developers in respiratory and other diseases. It focuses on where to find information about consumer values and preferences, at what points in the guideline development process to integrate their values and preferences, and why.**

***Methods*: In this review, we addressed the following questions: (*1*) What do we mean by "consumers"? (*2*) Why integrate the values and preferences of consumers of healthcare (patients, caregivers, and the public) into clinical practice guidelines? (*3*) What are the sources of information on consumer values? (*4*) When and how should consumer values and preferences be integrated into chronic obstructive pulmonary disease guideline recommendations?**

**We defined consumers as patients, caregivers, and members of the public, excluding groups that may also be identified as consumers of guidelines including health professionals, providers, and commissioners of services. We searched PubMed and other databases of methodological studies for existing systematic reviews and relevant methodological research. We did not conduct systematic reviews ourselves. Our conclusions are based on available evidence, supplemented by a rapid appraisal of a selection of qualitative studies, experience of what guideline developers are doing, and workshop discussions.**

***Results*: A clear distinction needs to be made between the use of information on consumer values and preferences by guideline developers, and the direct involvement of consumers in guideline development processes. Sources of information on consumer values include the research literature and direct elicitation of values both from organizations representing consumer interests and from individuals. To complement the identification of consumer values, there are a range of methods for involving consumers at all stages of guideline development, from consultation to direct membership of guideline development groups.**

***Conclusions*: Evidence-based guidelines need to consider explicitly the values and preferences of all relevant stakeholders (including those of consumers) and to provide opportunities for patients, caregivers, and the public to engage in the processes that consider and integrate those values into the development of guideline recommendations.**

## INTRODUCTION

Professional societies and other producers of evidence-based guidelines have recognized the need to involve a broad range of stakeholders in the development of guidelines, including consumers of healthcare (patients, caregivers, and the public) (1–4). In June 2007, the American Thoracic Society (ATS) and the European Respiratory Society (ERS) convened an international workshop of methodologists and researchers from around the world to coordinate efforts in guideline development using chronic obstructive pulmonary disease (COPD) as a model (5). Participants completed the work during the subsequent 4 years to develop a series of recommendations. This is the eighth of a series of 14 articles that were prepared to advise guideline developers in respiratory and other diseases on approaches for guideline development. This article focuses on how to involve consumers and integrate their values and preferences in guideline development.

## METHODS

The authors of this article developed and discussed several key questions (Table 1) and used them to update a review of the literature addressing the integration of consumer values and preferences in guideline development (6). In the absence of any systematic reviews of studies addressing consumer values relating to COPD, we searched PubMed and other databases for individual research studies that focused on COPD. These searches were supplemented by a rapid appraisal of 34 qualitative studies concerning consumer involvement in guidelines undertaken by the German Institute for Quality and Efficiency in Health Care (IQWiG) (personal communication). Our results and conclusions are also based on available evidence from the published and gray literature, from authors with experience in guideline development and consumer involvement methodologies, and workshop discussions.

## RESULTS

### What Do We Mean by "Consumers"?

People with a legitimate interest in guideline development include those who deliver care (healthcare professionals), those who finance, commission, manage, and assure care (government, policymakers, healthcare providers, and regulators), those who develop and manufacture healthcare products, and those who receive care (consumers of healthcare) (7). This paper focuses on the latter group and considers why, when, and how to ensure that their values and preferences are integrated into clinical practice guidelines, with a specific focus on COPD.

---

This article is a section of "Integrating and Coordinating Efforts in Chronic Obstructive Pulmonary Disease (COPD) Guideline Development," an American Thoracic Society (ATS) and European Respiratory Society (ERS) Workshop Report. This official ATS/ERS Workshop Report was adopted by the ATS Board of Directors, August 2012, and by the ERS Executive Committee, February 2012.

Funding for this conference was made possible (in part) by 1R13HL 90485-01 from the National Institutes of Health. The views expressed in written conference materials or publications and by speakers and moderators do not necessarily reflect the official policies of the Department of Health and Human Services; nor does mention of trade names, commercial practices, or organizations imply endorsement by the U.S. Government.

H.J.S. is supported by a European Commission: The human factor, mobility and Marie Curie Actions Scientist Reintegration Grant (IGR 42192) "GRADE."

**Proc Am Thorac Soc Vol 9, Iss. 5, pp 262–268, Dec 15, 2012**

**DOI: 10.1513/pats.201208-061ST**
**Internet address: www.atsjournals.org**

**TABLE 1. QUESTIONS ADDRESSED REGARDING CONSUMER INVOLVEMENT AND THE INTEGRATION OF CONSUMER VALUES IN GUIDELINE DEVELOPMENT**

1. What do we mean by "consumers"?
2. Why integrate the values and preferences of consumers of healthcare (patients, caregivers, and the public) into clinical practice guidelines?
3. What are the sources of information on consumer (patient, caregiver, and public) values?
4. When and how should consumer (patient, caregiver, and public) values and preferences be integrated into COPD guideline recommendations?

*Definition of abbreviation*: COPD = chronic obstructive pulmonary disease.

Consumers of healthcare include: (*1*) individual patients; (*2*) caregivers, including patients' family and friends (i.e., those who provide nonreimbursed care and/or support to patients); (*3*) members of the public (both as potential patients and as funders of healthcare through taxation, insurance, or direct payments); (*4*) voluntary and community organizations that represent the interests of patients, caregivers and the public; and (*5*) advocates representing the interests of patients, caregivers, and other client groups. They are described collectively as "consumers" (without implying consumerist assumptions about health services) and are distinct from other consumers of guidelines, such as health professionals, commissioners, and providers of services.

## Why Integrate Consumer Values and Preferences into the Formulation of Clinical Practice Guidelines?

In this section, we consider the integration of consumer values and preferences by describing separately those of patients and their nonprofessional caregivers, and those of the general public.

***Patient and caregiver values and preferences.*** Whereas early descriptions of evidence-based medicine focused on bringing evidence to bear on care of the individual patient (8), more recent expositions have evolved in two major ways. First, they have emphasized that values and preferences are essential for clinical decision making and, thus, research evidence that does not include evidence about values and preferences, in and of itself, is never sufficient to fully inform clinical decision making (9). Second, they have extended the logic of incorporating evidence and patients' values into individual clinical decisions, to incorporating evidence and patients' values into guideline development (10).

To be fully integrated into clinical guidelines, patient values and preferences must inform both the guideline development process (including what evidence is brought to bear on the recommendations) and the tradeoffs involved in making recommendations. The ultimate aim is for guidelines to address the issues that are important to patients and caregivers, and to be sensitive to the range of values and preferences held by patients and caregivers. Patients have a unique perspective on their condition, on what constitutes good and poor care, and on the outcomes they hope to achieve (and avoid) as a result of any intervention (11). The rationale for inclusion of these values can be conceptualized in three ways: integration of the healthcare experience to improve the quality of guidelines, increased legitimacy for the guidelines if the process is more open and transparent, and the fundamental principle that patients are affected by decisions and should have an opportunity to provide input (12).

Caregiver or proxy values are important in two ways. First, they are important in articulating the values and preferences of patients who may be unable to speak for themselves (for example, patients who are critically ill). Second, they are important in highlighting the physical, emotional, and financial needs of the caregivers who provide practical and emotional support to patients. Some patients will choose to have decisions based on their family members' values and preference even when they have decisional capacity. This perspective is more common in some cultures, and guidelines on some topics may need to address this issue (13). However, it is important to recognize that caregivers' and patients' interests sometimes conflict (14). Furthermore, family members, nonprofessional caregivers, and health professionals all share large degrees of inaccuracy in ascertaining or predicting patient wishes or expectations (15).

The published research available to a guideline development group may not have taken into account the range of outcomes that patients identify as important, or considered the range of interventions that may achieve those outcomes. For instance, COPD research studies may focus on lung function, which is less relevant to patients than symptoms, quality of life, and functional status.

Research focused on surrogate outcomes may, in turn, drive the priorities of guideline developers, sometimes at the expense of outcomes or interventions that patients consider important. A systematic review found that randomized controlled trials often overlooked evidence of effectiveness of innovations in nurse-led chronic disease management for patients with COPD; patient issues including self-management skills, coping, or self confidence; patients' and caregivers' satisfaction with the interventions or their preferences for care; and caregivers' quality of life (16). That patients and caregivers consider quality of life issues important (17–19) argues strongly for their integration into guideline recommendations.

Addressing patients' and caregivers' values and preferences may help make the guideline recommendations more acceptable to them (20) and, thus, their implementation more likely. A range of factors may influence how patients (and caregivers) perceive the intervention. These include patient and caregiver knowledge, attitudes, expectations, and beliefs; patient goals for life and health; prior experience with the intervention and the condition; symptom experience (for example, breathlessness, pain, dyspnea, and weight loss); preferences for desirable and undesirable outcomes; perceived impact of the condition or interventions on physical and social functioning, daily activities, work, sports, and leisure activities; psychosocial issues (quality of life, well-being, satisfaction, depression, anxiety, grief, and loss); interactions between the work of implementing the intervention, the intervention itself, and other work and contexts the patient may be experiencing; preferences for alternative courses of action; and preferences relating to communication content and styles, information, and involvement in decision making and care (including self-management).

***The values of the general public.*** Involving the wider general public involves challenges beyond involvement of COPD patients and their caregivers. Public views are often based on hypothetical judgments (compared with patients and caregivers, who can draw on personal experience). Public views on the use of public funds, for instance, may differ from those of people affected by the condition (for example, placing a higher value on life-prolonging treatment vs. improvement in quality of life) (21).

The general public may have beliefs that conflict with the interests of patients and caregivers. This may be critically important for guideline panels who may be increasingly expected to take resource use into account (22). For example, the public may not be aware of the importance of COPD (23) and may be unwilling to recommend resource-intensive treatments for

conditions that are perceived as "self-inflicted": the stigma associated with smoking adversely affects people with lung cancer (24) and COPD (25). The wider public also includes patients with conditions other than COPD and, therefore, some of them may have competing interests where health service resources are limited.

### What Are the Sources of Information on Consumer Values?

Sources of information on consumer values include the published literature and direct elicitation.

***The published literature.*** Although we found no systematic review of studies addressing consumer values relating to COPD, we did find a systematic review of patients' experiences of breathlessness, which is largely about COPD (26). Individual studies, both qualitative and quantitative in nature, have explored both the outcomes patients and caregivers value and patient views on or preferences for specific interventions.

Studies of outcomes that patients consider important often reveal an interrelationship between positive and negative consequences of interventions. For example, oxygen use may prolong life but result in social isolation arising from both the physical and social restrictions it imposes and embarrassment associated with its use (27). The ways in which patients evaluate breathlessness may be influenced by its physical impact (lack of mobility and being housebound [28]), but also by the extent to which patients experience anxiety, panic, fear, frustration, and tiredness (29, 30). Studies have explored the impact of COPD and exacerbations on the symptoms and daily life of patients and caregivers (31, 32), but have also noted the influence on outcomes of patients' understanding of the term "exacerbation" (which may be underestimated by physicians) (33), how patients interpret symptoms and warning signs of exacerbations (34), and the need to know more about what triggers exacerbations and how to self-manage (35).

Studies of patient preferences for specific interventions have explored issues ranging from complementary and alternative medicines (36) to smoking cessation (37) and resuscitation (38). Researchers have also made attempts to quantify the impact of disease on daily life and well-being in a formal and standardized manner, resulting in the development of three distinct types of instruments to measure "health-related quality of life" (39). First, utility scales attempt to quantify different states of health on a continuum from perfect health to death, an approach that has been particularly favored by health economists. Second, general or generic health measures aim to provide valid estimates of impaired health in chronic respiratory disease and include the Sickness Impact Profile, the Short-Form 36-item questionnaire (SF-36), and the Nottingham Health Profile. Third, disease-specific instruments (such as the Chronic Respiratory Questionnaire or the St. George's Respiratory Questionnaire) aim to be more relevant to an individual's health problems and more sensitive to small changes in response to therapeutic interventions (40).

The expansion of treatment options for COPD has resulted in increasing numbers of studies that collect pharmacoeconomic evidence in an effort to complement treatment guidelines. However, a review of pharmacoeconomic studies on drug therapy for stable COPD concluded that different comparators prevent direct comparison of studies and prevent a comprehensive assessment of any particular therapy as being cost-effective across settings and different patient populations (41). Furthermore, patient groups have voiced concerns about attempts that are seen as reducing patient values to numerical scores that neither include nor give sufficient weight to the range of dimensions of quality of life and the experience of treatment, particularly in diverse contexts that patients consider and see as potentially discriminating against certain groups including older people (42). The specific measures used depend on the research question and population under study, but it is often advisable to include a generic and a disease-specific measure, and to consider a utility measure if cost analyses are anticipated.

***Direct elicitation of consumer values.*** Guideline developers have directly elicited consumer values and preferences using a range of methods including membership of the guideline panel; separate patient panels whose views inform the main panel; elicitation of patient values through workshops, focus groups, or interviews; and consultation on guideline products (43–45).

The World Health Organization (WHO) has made global recommendations about involving end users, and patients specifically, as members of guideline panels (46). The Appraisal of Guidelines for Research and Evaluation (AGREE) instrument includes patient involvement as a principle criterion in assessing the quality of guidelines (47). In 1996, Bastian proposed that consumers had an important role and vested interest in deciding what *guidelines* address, how they will be developed, and what they will say (48). The National Institute for Health and Clinical Excellence (NICE) suggests that a key role for patient and caregiver members is to ensure that the views, experiences, and interests of patients inform the panel's work (49). The Institute of Medicine report on transparent guideline development suggests the involvement of patients or patient representatives during the guideline development process (50).

Some clinical guideline programs have involved either individual patients and caregivers or people from organizations representing the interests of patients in the development of both nationally and locally produced guidelines. A review published in 2000 found that only 25% of organizations regularly did so (51). A 2003 review of five national and two international COPD guidelines found "little or no consumer input" (52). Table 2 shows how patients have been involved in COPD

**TABLE 2. CONSUMER INVOLVEMENT IN COPD GUIDELINES**

| | Consultation | Patient GDG Members | Patient Guides/Decision Aids |
|---|---|---|---|
| ACCP, USA | | ✓ | |
| AQuMed | With patients on patient guide | | Patient guide |
| Australia/New Zealand | | ✓ | |
| CBO, The Netherlands | Quote questionnaire<br>Focus groups<br>Patient panel | ✓ | Patient guide |
| IQWIG, Germany | Consumer committee | | |
| GOLD, International | | | Decision aids |
| NICE, England and Wales | With patient and caregiver organizations | ✓ | Patient version of guideline |

*Definition of abbreviations*: ACCP = American College of Chest Physicians; CBO = Centraal BegeleidingsOrgaan; COPD = chronic obstructive pulmonary disease; GDG = Guideline Development Group; GOLD = Global Initiative for Chronic Obstructive Lung Disease; IQWIG = Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen; NICE = National Institute for Health and Clinical Excellence.

guidelines published since that review. The table also highlights guidelines where information about the recommendations is available in patient-friendly formats, either in the form of patient guides or decision aids.

Eliciting information has generally focused on patients and caregivers, but not on the wider public. The NICE Citizens Council is an example of an attempt to integrate the values of members of the public without a condition-specific interest. It is composed of 30 members of the public representing the sociodemographic characteristics of England and Wales, and provides overarching principles for guideline development, regardless of the topic (53). The council considers ethical questions, such as whether NICE guidance should give special consideration to certain patient groups, depending on factors such as age, illnesses perceived by the public as self-inflicted, and groups where there may be health inequalities. Guideline developers are expected to take the "social value" judgments of the Council into account when making recommendations (54).

### When and How Should Consumer Values and Preferences Be Integrated into Guidelines?

Deciding when to integrate consumer values and preferences into guidelines involves consideration of the stages and activities of guideline development. Initial stages involve selecting the topic and determining the focus of the guideline (what issues the guideline will and will not address). The next stage relates to the activities of the guideline group in developing the clinical questions, searching for evidence, synthesizing and grading the quality of the evidence, and then generating, refining, and grading recommendations.

We discuss below when and how consumer values can feed into each of these stages, noting that some guideline development agencies have formally documented approaches to consumer involvement at such key stages (55–57). We also consider how to identify consumer contributors to the guideline development process.

***Topic selection.*** The Scottish Intercollegiate Guidelines Network and NICE both invite any group or individual to propose a guideline topic for future development. In addition, groups that consider and prioritize topics include consumer members, who help ensure that topic selection takes account of consumer values and is not solely determined by professional or other (e.g., pharmaceutical industry) interests.

***Determining the focus and boundaries of the guideline.*** The parameters given to guideline developers will determine the boundaries for setting clinical questions and subsequently the breadth of the evidence search and resulting recommendations. If consumer values are not integrated into this stage of the process, the guideline may not make recommendations about issues of particular concern to them. Indeed, there is considerable variation in the extent to which COPD guidelines make recommendations about some key issues of importance to consumers, including self management, enhancing quality of life, psychosocial issues, caregiver impact, patient involvement in the therapeutic relationship, and meeting consumer information needs (H. Bastian, personal communication).

***The work of the guideline development group.*** The work plan of a guideline development group provides opportunities for integrating consumer values into the clinical questions, the appraisal of the evidence, and the development of recommendations. Involving consumers directly in the process may help overcome some of the deficiencies of current COPD guidelines. For example, patients with COPD are less likely to receive care that is consistent with their values for end-of-life care than patients with lung cancer (58) or cancer in general (59). Successful integration of consumer values should include consideration of variations in values between different subgroups. For example, different recommendations may be required for people who are willing and able to give up smoking than for people who do not stop smoking.

***Commenting on draft recommendations.*** Consultation with consumers on draft recommendations helps ensure that their values have been integrated into the recommendations. They can provide feedback regarding whether the recommendations are consistent with the range of their values and preferences (taking account of desired outcomes, the ways in which people weigh up risks and benefits, and preferred treatment and management options), and whether they are practical in the "real" world.

***Identifying consumers able to contribute.*** A further challenge lies in identifying consumers who are willing and able to contribute directly to guideline development. Professional resistance to consumer membership of guideline panels sometimes reflects concerns about the possible idiosyncratic or unrepresentative perspectives of consumer participants. However, this can apply to all participants, not just consumers. Clarity about the roles of consumer members, open and transparent selection processes, and training and support processes can address some of these concerns and maximize consumer contributions.

Clarity about roles is important because consumer members of panels are sometimes criticized if they are too knowledgeable and, therefore, seen as divorced from grassroots experiences. Ironically, they may also be criticized if their lack of experience or research knowledge inhibits their useful participation in discussions. However, one should recognize that consumer members are not expected to provide definitive answers, but to raise the questions that will help ensure that consumer values and preferences are adequately considered.

Guideline developers should consider that the pace of the process can challenge meaningful direct consumer involvement. Contributing factors include limited time to come to terms with the concept of evidence-informed guidelines, the practicalities of involving people, and the fact that some people may not be able to see the development process through from start to finish due to ill health and poor prognosis.

NICE and others have argued that some of these issues can be addressed by having open, transparent, and clear processes that make explicit the roles of consumer participants and the opportunities (and boundaries) for participation at different stages of guideline development. For people who actively contribute to guideline development as consumer members of guideline development groups, this includes advertising vacancies, providing job descriptions that detail the task to be undertaken, and selecting applicants according to explicit criteria based on background and experience needed (55). Preparation, training, and support can facilitate useful productive consumer participation (60) and is advocated by guideline panel chairs and consumer members with experience of such provision (61).

Finally, it is also important that consumer representatives disclose potential conflicts of interest in guideline development and that these conflicts of interest are addressed in a similar fashion to those of others involved in the guideline development process.

## DISCUSSION

In this review, we identified literature suggesting a number of reasons for ensuring both that consumer values are integrated into guideline development and that consumers are involved in the guideline development process. The underlying rationale is that consumers have a unique perspective on their condition, on what constitutes good and poor care, and on the outcomes they hope to achieve (and avoid) as a result of any intervention

**TABLE 3. SUGGESTIONS FOR INTEGRATING CONSUMER VALUES AND PREFERENCES INTO THE DEVELOPMENT OF COPD GUIDELINES**

1. Guideline developers should make an explicit commitment to integrate consumer values into guideline development processes and products. The goal is to take account of consumer views and preferences and, where possible, accommodate different consumer preferences.
2. A review of the literature on consumer values and preferences should be an integral part of the guideline development process. Identified values and preferences (relating, e.g., to interventions, comparators, and outcomes considered by the guideline group) can inform the refining of clinical questions, the interpretation of evidence from the research literature, and the development and wording of recommendations.
3. Guideline developers should consider involving consumers in the guideline development process, indirectly and/or directly. Indirect involvement includes consultation with representative consumer groups and/or surveys or focus groups with guideline-relevant consumers to obtain information on their values and preferences. Such information may then inform the deliberations of the guideline development group. Direct involvement typically involves recruiting consumers as members of guideline steering, development, and/or working groups. Such involvement helps ensure that consumers have the opportunity to influence, alongside professional members, both the deliberations and outputs of the guideline development group.
4. Guideline developers should assess the training needs of both professional and consumer participants and provide adequate support and training to promote effective collaborative working.
5. Further research and evaluation is needed to establish the most effective ways in which consumers values and preferences can be incorporated into clinical guidelines.

*Definition of abbreviation*: COPD = chronic obstructive pulmonary disease.

(11). The rationale for inclusion of these values can be conceptualized in three ways: integration of healthcare experience to improve the quality of guidelines, increased legitimacy for the guidelines if the process is more open and transparent, and the fundamental principle that patients are affected by decisions and should have an opportunity to provide input (12).

Integrating consumer values into guideline recommendations presents guideline developers with a number of challenges. High-quality studies addressing consumer values and preferences may be difficult to locate or may not exist. When such information can be obtained, either from the research literature or through direct elicitation, there is variability in patient values and preferences at different stages of disease (new-onset, stable, acute exacerbation, rehabilitation, palliation, and terminal care), with different disease severity (from mostly asymptomatic to critically ill), and when considering different issues (for example, when considering testing, medication choice, surgical treatments, intubation, and whether to enter a hospice). Values and preferences may differ across age and sex, socioeconomic status, ethnicity, and culture. It remains important, however, to ensure that guideline statements are sensitive to heterogenous preferences, especially when a potential recommendation has pros and cons (for example, involving the need to balance health gain with adverse side effects).

A clear distinction needs to made between the use of information on values and preferences (whether obtained from the literature or by inviting contributions directly) for others to consider and the direct involvement of patients, caregivers, or the public in the decision-making process. Restricting involvement to the former approach is at odds with WHO recommendations (46) and with national and international health policy initiatives that advocate and foster opportunities for patient and public participation at all stages of decision making relating to health policies, strategies, and medical research (62–67). However, restricting involvement to the latter approach can raise concerns about identifying participants who can represent the range of consumer experience and may eliminate input from vulnerable or marginalized populations for which some research literature may exist.

Past initiatives to involve consumers directly in guideline development have revealed that involvement can be resource intensive and consumers' contributions are sometimes limited (43). Challenges include identifying and/or supporting consumers who are willing and able to contribute directly to guideline development. Although open recruitment processes, preparation, training, and support can facilitate consumer participation and are advocated by guideline panel chairs and patient/caregiver members with experience (59), it has been argued that many professional societies or small healthcare agencies do not have the resources to provide this level of support (68).

Recognition that consumer values and preferences may differ significantly from those of clinical experts is the first step in enhancing the relevance and pertinence of recommendations. This recognition may also challenge guideline methodologists and guideline producers to explore strategies that will achieve true evidence-based guidelines. Such guidelines will not only explicitly consider the quality and findings of the best available research evidence, but also take into account the values and preferences of all relevant stakeholders, most critical among these, the values and preferences of patients, caregivers, and the public. To this end, we make a number of suggestions for addressing the issue of integrating consumer values and preferences into guideline development (Table 3).

Although we recognize that not all developers will have the resources to commit to all these suggestions, we hope that the list may help developers identify and prioritize areas for initiating or developing activity in this area.

***Author Disclosures***: M.K., E.A.A., and H.B. reported no commercial interests relevant to subject matter. F.C. consulted with and received lecture fees from NICE (National Institute for Clinical Excellence, U.K.). J.R.C. reported no commercial interests relevant to subject matter. G.G. was a member of the GRADE Working Group. V.M.M. and S.O. reported no commercial interests relevant to subject matter. H.J.S. was a member of the GRADE Working Group, which received honoraria and lecture fees regarding GRADE. All workshop participants disclosed receipt of honoraria and travel reimbursement from the American Thoracic Society.

## References

1. Brouwers MC, Kho ME, Browman GP, Burgers J, Cluzeau F, Feder G, Fervers B, Graham ID, Hanna SE, Makarski J; AGREE Next Steps Consortium. Validity assessment of items and tools to support application: development steps towards the AGREE II – Part 2. *Can Med Assoc J* 2010;182:E472–E478 10.1503/cmaj.091716.
2. Kelson M. Patient involvement in clinical guideline development – where are we now? *J Clin Gov* 2001;9:1.
3. Kelson M. NICE's commitment to patient, caregiver and public involvement. In: Littlejohns P, Rawlins M, editors. Patients, the public and priorities in healthcare. Oxford: Radcliffe Publishing Ltd; 2009. pp. 9–18.
4. Oxman A, Freitheim A, Schünemann H; SURE. Improving the use of research evidence in guideline development: introduction. *Health Res Policy Syst* 2006;4:12.
5. Schunemann HJ, Woodhead M, Anzueto A, Buist S, Macnee W, Rabe KF, Heffner J. A vision statement on guideline development for respiratory disease: the example of COPD. *Lancet* 2009;373:774–779.
6. Schunemann HJ, Freitham A, Oxman D; Improving the use of research evidence in guideline development: 10. Integrating values and consumer involvement. *Health Res Policy Syst* 2006;4:22.
7. Cluzeau F, Wedzicha J, Kelson M, Corn J, Kunz R, Walsh J, Schünemann H. Integrating values and consumer involvement in guidelines. Stakeholder

involvement: how to do it right. Article 9 in integrating and coordinating efforts in COPD guideline development. An official ATS Workshop Report. *Proc Am Thorac Soc* 2012;9:269–273.
8. Evidence-Based Medicine Working Group. Evidence-based medicine: a new approach to teaching the practice of medicine. *JAMA* 1992;268:2420–2425.
9. Guyatt GH, Haynes RB, Jaeschke RZ, Cook DJ, Green L, Naylor CD, Wilson MC, Richardson WS; Evidence-Based Medicine Working Group. Users' guides to the medical literature: XXV. Evidence-based medicine: principles for applying the users' guides to patient care. *JAMA* 2000;284:1290–1296.
10. Eddy DM. A manual for assessing health practices and designing practice policies: the explicit approach. Philadelphia: American College of Physicians; 1992.
11. Kelson M. Patient defined outcomes. London: College of Health; 1999.
12. Van de Bovenkamp HM, Trappenburg MJ. Reconsidering patient participation in guideline development. *Health Care Anal* 2009;17:198–216.
13. Blackhall L, Murphy ST, Frank G, Michel V, Azen S. Ethnicity and attitudes toward patient autonomy. *JAMA* 1995;274:820–825.
14. Entwistle V, Renfrew M, Yearley S, Forrester J, Lamont T. Lay perspectives: advantages for health research. *BMJ* 1998;316:463–466.
15. Shalowitz DI, Garrett-Mayer E, Wendler D. The accuracy of surrogate decision makers: a systematic review. *Arch Intern Med* 2006;166:493–497.
16. Taylor S, Candy B, Bryar R, Ramsay J, Vrijhoef H, Esmond G, Jadwiga A, Wedzicha J, Griffiths J. Effectiveness of innovations in nurse led chronic disease management for patients with chronic obstructive pulmonary disease: systematic review of evidence. *BMJ* 2005;331:485.
17. Jones P. Quality of life measures for patients with diseases of the airways. *Thorax* 1991;46:676–682.
18. Van der Molen T, Pieters W, Bellamy D, Taylor R. Measuring the success of treatment for chronic obstructive pulmonary disease – patient, physician and healthcare payer perspectives. *Respir Med* 2002;96:S17–S21.
19. Elkington H, White P, Addington-Hall J, Higgs R, Pettinari C. The last year of life of COPD: a qualitative study of symptoms and services. *Respir Med* 2004;98:439–445.
20. Iscoe N. We may get this horse to water, but will it drink? *Can Med Assoc J* 1998;158:345–346.
21. Bowling A, Jacobson B, Southgate L. Explorations in consultation of the public and health professionals on priority setting in an inner London health district. *Soc Sci Med* 1993;37:851–857.
22. Hill OL Sr, Falck-Ytter Y, Cruz A, Atkins D, Baumann M, Jaeschke R, Woitalla T, Schünemann HJ. Incorporating considerations of cost-effectiveness, affordability, and resource implications in guideline development. Article 6 in integrating and coordinating efforts in COPD guideline development. An official ATS Workshop Report. *Proc Am Thorac Soc* 2012;9:251–255.
23. Kesten S, Ulrich J, Rich D. Public awareness of chronic obstructive pulmonary disease [abstract]. *Chest* 2003;124:167S.
24. Chapple A, Ziebland S, McPherson A. Stigma, shame, and blame experienced by patients with lung cancer: qualitative study. *BMJ* 2004;328:1470.
25. Knauft E, Nielsen EL, Engelberg RA, Randall J, Curtis J. Barriers and facilitators to communication about end-of-life care for patients with COPD. *Chest* 2005;127:2188–2196.
26. Gysles M, Bausewein C, Higginson IJ. Experience of breathlessness: a systematic review of the qualitative literature. *Palliat Support Care* 2007;5:281–302.
27. Earnest M. Explaining adherence to supplemental oxygen therapy: the patient's perspective. *J Gen Intern Med* 2002;17:749–755.
28. Elkington H, White P, Addington-Hall J, Higgs R, Pettinari C. The last year of life of COPD: a qualitative study of symptoms and services. *Respir Med* 2004;98:439–445.
29. Barnett M. Chronic obstructive pulmonary disease: a phenomenological study of patients experiences. *J Clin Nurs* 2005;14:805–812.
30. Bailey P. The dyspnea-anxiety-dyspnea cycle – COPD patients' stories of breathlessness: "It's scary / when you can't breathe. *Qual Health Res* 2004;14:760–778.
31. Miravitles M, Anzueto A, Legnani D, Forstmeier L, Fargel M. Patient's perception of exacerbations of COPD: the PERCEIVE study. *Respir Med* 2007:101:453–460.
32. Seamark D, Blake D, Seamark C, Halpin D. Living with severe chronic obstructive pulmonary disease (COPD): perceptions of patients and their caregivers. An interpretative phenomenological analysis. *Palliat Med* 2004;18:619–625.
33. Kessler R, Stahl E, Vogelmeier C, Haughney J, Trudeau E, Lofdahl C, Partridge M. Patient understanding, detection and experience of COPD exacerbations: an observational, interview-based study. *Chest* 2006;130:133–142.
34. Small S, Graydon J. Uncertainty in hospitalized patients with chronic obstructive pulmonary disease. *Int J Nurs Stud* 1993;30:239–246.
35. Monninkhof E, van der Aa M, van der Valk P, van der Palen J, Zielhuis G, Koning K, Pieterse M. A qualitative evaluation of a comprehensive self-management programme for COPD patients: effectiveness from the patients' perspective. *Patient Educ Couns* 2004;55:177–184.
36. George J, Iaonnides-Demos LL, Santamaria NM, Kong DC, Stewart K. Use of complementary and alternative medicines by patients with chronic obstructive pulmonary disease. *Med J Aust* 2004;181:248–251.
37. Robinson T. Living with severe hypoxic COPD: the patient's experience. *Nurs Times* 2005;101:38–42.
38. Gaber KA, Barnett M, Planchant Y, McGavin CR. Attitudes of 100 patients with chronic obstructive pulmonary disease to artificial ventilation and cardio-pulmonary resuscitation. *Palliat Med* 2004;18:626–629.
39. Guyatt GH, Feeny DH, Patrick DL. Measuring heath-related quality of life. *Ann Intern Med* 1993;118:622–629.
40. Mahler D. How should health-related quality of life be assessed in patients with COPD? *Chest* 2000;117:54S–57S.
41. D'Souza A, Smith M, Miller L, Kavookjian J. An appraisal of pharmacoeconomic evidence of maintenance therapy for COPD. *Chest* 2006; 129:1693–1708.
42. House of Commons Health Select Committee. National Institute for Health and Clinical Excellence. First report of session 2007–08. London: The Stationery Office Limited; 2008. HC 27-I.
43. Van Wersch A, Eccles M. Involvement of consumers in the development of evidence based clinical guidelines: practical experiences from the North of England evidence based guideline development programme. *Qual Health Care* 2001;10:10–16.
44. Kelson M. The NICE Patient Involvement Unit. *Evid Based Healthc Public Health* 2005;9:304–307.
45. Kelson M; World Health Organization. Patient and caregiver membership of NICE clinical guideline development groups. *Essent Drugs Monit* 2005;34:21.
46. Global Programme on Evidence for Health Policy, World Health Organization. Guidelines for WHO guidelines. Geneva: World Health Organization; 2003. EIP/GPE/EQC/2003.1.
47. Cluzeau FA, Burgers JS, Brouwers M, Grol R, Mäkelä M, Littlejohns P, Grimshaw J, Hunt C; The AGREE Collaboration. Development and validation of an international appraisal instrument for assessing the quality of clinical practice guidelines: the AGREE project. *Qual Saf Health Care* 2003;12:18–23.
48. Bastian H. Raising the standard: practice guidelines and consumer participation. *Int J Qual Health Care* 1996;8:485–490.
49. National Institute for Health and Clinical Excellence. Contributing to a NICE clinical guideline: a guide for patients and caregivers. London: NICE; 2006.
50. Institute of Medicine. Current best practices and standards for development of trustworthy CPGS: Part II, Traversing the process. In: Graham R, Mancher M, Miller Wolman D, Greenfield S, Steinberg E, editors. Clinical practice guidelines we can trust. Washington, D.C.: The National Academies Press; 2011.
51. Health Policy Forum. The 2000 Consensus Conference on Clinical Practice Guidelines. November 2nd Toronto; 2000.
52. Smith BJ, Hender K, Frith P, Crockett AJ, Cheok F, Spedding S. Systematic assessment of clinical practice guidelines for the management of chronic obstructive pulmonary disease. *Respir Med* 2003;97:37–45.
53. Littlejohns P, Rawlins M, editors. Patients, the public and priorities in healthcare. Oxford: Radcliffe Publishing Ltd; 2009.
54. National Institute for Health and Clinical Excellence. Social value judgment: principles for the development of NICE guidance. London: NICE; 2005.
55. Scottish Intercollegiate Guidelines Network. SIGN 50: a guideline developers' handbook. Section 8: Consultation and peer review [revised 2008 Jan; accessed 2011 Jun 20]. Available from: http://www.sign.ac.uk/guidelines/fulltext/50/section3.html
56. The National Institute for Health and Clinical Excellence. The guideline development process: an overview for stakeholders, the public and the NHS, 4th ed. London: NICE; 2008 (Accessed 20 June 2011). Available

from: http://www.nice.org.uk/aboutnice/howwework/developingniceclinicalguidelines/clinicalguidelinedevelopmentmethods/clinical_guideline_development_methods.jsp?domedia=1&mid=62F02D9B-19B9-E0B5-D4A26EC9A934FDC7
57. Thomas V. Patient and caregiver involvement in NICE clinical guidelines. In: Littlejohns P, Rawlins M, editors. Patients, the public and priorities in healthcare. Oxford: Radcliffe Publishing Ltd; 2009. pp. 19–27.
58. Claessens MT, Lynn J, Zhong Z, Desbiens NA, Phillips RS, Wu AW, Harrell FE Jr, Connors AF Jr. Dying with lung cancer or chronic obstructive pulmonary disease: insights from SUPPORT. *J Am Geriatr Soc* 2000;48:S146–S153.
59. Exley C, Field D, Jones L, Stokes T. Palliative care in the community for cancer and end-stage cardio-respiratory disease: the views of patients, lay-caregivers and health care professionals. *Palliat Med* 2005;19:76–83.
60. McIver S. Preparation and training: the key to better patient involvement. *Qual Health Care* 2001;10:3.
61. National Institute for Health and Clinical Excellence. A report on a study to evaluate patient/caregiver membership of the first NICE Guideline Development Groups [accessed 20 June 2011]. London: NICE; 2004. Available from: www.nice.org.uk
62. Department of Health. Research governance framework for health and social care. London: Department of Health; 2005.
63. National Health and Medical Research Council and Consumers' Health Forum of Australia. Statement on consumer and community participation in Health and Medical Research. Canberra: AusInfo; 2001.
64. National Institutes of Health Director's Council of Public Representatives [accessed 20 June 2011]. Bethesda, MD: National Institutes of Health. Available from: www.copr.nih.gov
65. Minister of Health. The New Zealand health strategy. Wellington, New Zealand: Ministry of Health; 2000.
66. The Cochrane Collaboration Consumer Network [accessed 15 November 2012]. Available from: http://www.consumers.cochrane.org/
67. Kelson M. Consumer collaboration, patient-defined outcomes and the preparation of Cochrane Reviews. *Health Expect* 1999;2:129–135.
68. Culyer A. Involving stakeholders in healthcare decisions: the experience of the National Institute for Clinical Excellence. *Healthc Q* 2005;8:56–60.

# Stakeholder Involvement: How to Do It Right

## Article 9 in Integrating and Coordinating Efforts in COPD Guideline Development. An Official ATS/ERS Workshop Report

**Françoise Cluzeau, Jadwiga A. Wedzicha, Marcia Kelson, Judy Corn, Regina Kunz, John Walsh, and Holger Schünemann; on behalf of the ATS/ERS Ad Hoc Committee on Integrating and Coordinating Efforts in COPD Guideline Development**

***Introduction*: Professional societies, like many other organizations around the world, have recognized the need to use more rigorous processes to ensure that healthcare recommendations are informed by the best available research evidence with input from appropriate stakeholders. This is the ninth of a series of 14 articles that were prepared by an international panel to advise guideline developers in respiratory and other diseases on approaches for guideline development. We updated a review of the literature on stakeholder involvement, focusing on six key questions.**

***Methods*: In this review we addressed the following questions. (*1*) What are "stakeholders"? (*2*) Why involve stakeholders in guidelines? (*3*) At what stage should stakeholders contribute to guidelines? (*4*) What are the potential barriers to integrating stakeholder involvement? (*5*) How can stakeholders be involved effectively? (*6*) Should anyone be excluded from the process? We searched PubMed and other databases of methodological studies for existing systematic reviews and relevant methodological research. We did not conduct our own systematic reviews. Our conclusions are based on available evidence, the experience of guideline developers, and workshop discussions.**

***Results and Discussion*: Stakeholders are all those who have a legitimate interest in a guideline. They include healthcare professionals, patients and caregivers, public and private funding bodies, managers, employers, and manufacturers. Their engagement is justified for several reasons, including limitations of evidence, principles of transparency and democracy, ownership, and potential policy implications. They have a role to play at different points of guideline development, but their involvement can be complex. To be successful, stakeholder engagement needs to be inclusive, equitable, and adequately resourced.**

## INTRODUCTION

Professional societies, like many other organizations around the world, have recognized the need to use more rigorous processes to ensure that healthcare recommendations are informed by the best available research evidence. Engaging the participation of stakeholders is an accepted feature of high-quality clinical guideline development (1). In June 2007, the American Thoracic Society (ATS) and the European Respiratory Society (ERS) convened an international workshop of methodologists and researchers from around the world for coordinating efforts in guideline development using chronic obstructive pulmonary disease (COPD) as a model. This is the ninth of a series of 14 articles that were prepared by an international panel to advise guideline developers in respiratory and other diseases on approaches for guideline development. This article focuses on the merits of involving stakeholders in guideline development and how this can be achieved effectively in the context of respiratory disease guidelines. This article complements two others published in this issue: one on patient and caregiver involvement (2) and one on guideline development and group processes (3).

## METHODS

The authors of this article developed and discussed the key questions in this article. We updated a review of the literature on stakeholder involvement, focusing on the six key questions in Table 1.

We searched PubMed and other databases of methodological studies for existing systematic reviews and relevant methodological research through 2010, using the search terms "stakeholders" or "users" or "public," for published literature on "stakeholder

### POTENTIAL STAKEHOLDER GROUPS FOR COPD GUIDELINES

- Care deliverers (professionals, managed care programs). They have to manage increasing consultations for COPD.
- Caregivers (they provide care to patients with COPD, often at home).
- Those receiving it (consumers or patients). COPD is disabling and restricts many everyday activities, such as walking up stairs.
- Those managing care (policy makers, public health services), having to plan new services for chronic COPD or new interventions for preventing COPD (such as stopping smoking).
- Those monitoring care (quality assurance companies). Ascertaining what strategies and treatments for managing COPD are effective.
- Those financing it (governments, health insurers, the public). They have to consider the costs incurred when exacerbations require hospital treatment.
- Employers. Time lost when patients with COPD do not work.
- Manufacturers, pharmaceutical companies, the healthcare industry who are interested in how to market their products. In COPD there are other relevant groups, such as oxygen delivery contractors.

This article is a section of "Integrating and Coordinating Efforts in Chronic Obstructive Pulmonary Disease (COPD) Guideline Development," an American Thoracic Society (ATS) and European Respiratory Society (ERS) Workshop Report. This official ATS/ERS Workshop Report was adopted by the ATS Board of Directors, August 2012, and by the ERS Executive Committee, February 2012.

Funding for this conference was made possible (in part) by 1R13HL90485–01 from the National Institutes of Health. The views expressed in written conference materials or publications and by speakers and moderators do not necessarily reflect the official policies of the Department of Health and Human Services; nor does mention of trade names, commercial practices, or organizations imply endorsement by the U.S. Government.

H.J.S. is supported by a European Commission: The human factor, mobility and Marie Curie Actions Scientist Reintegration Grant (IGR 42192) "GRADE."

**Proc Am Thorac Soc Vol 9, Iss. 5, pp 269–273, Dec 15, 2012**

**DOI: 10.1513/pats.201208-062ST**
**Internet address: www.atsjournals.org**

**TABLE 1. QUESTIONS ADDRESSED REGARDING STAKEHOLDER INVOLVEMENT IN GUIDELINE DEVELOPMENT**

1. What are "stakeholders"?
2. Why involve stakeholders in guidelines?
3. At what stage should stakeholders contribute to guidelines?
4. What are the potential barriers to integrating stakeholder involvement?
5. How can stakeholders be involved effectively?
6. Should anyone be excluded from the process?

involvement" and "clinical guidelines" in Medline and in the Cochrane library from 2005 to 2011, restricting the search to systematic reviews and reviews. We did not conduct systematic reviews ourselves. Our conclusions are based on available evidence from published literature, experience from guideline developers, and workshop discussions.

## RESULTS

### 1. What Are "Stakeholders"?

Broadly, stakeholders can be defined as people who have a legitimate interest in a guideline. Applying the example of COPD, which is likely to become the third most common cause of death worldwide by 2020, the burden of the disease is considerable from all relevant viewpoints (4, 5). The text box lists stakeholder groups relevant to a COPD guideline. All of these groups have an interest in a COPD guideline because they may affect and be affected by it at some point. They form a larger constituency than that involved in the "face-to-face" clinical setting or in consultation and treatment.

### 2. Why Involve Stakeholders in Guidelines?

Input from patients and caregivers is now widely promoted (6, 7), but involvement from the wider public is less well understood. Nevertheless, there is evidence that engaging relevant stakeholders has an impact on guideline uptake (8) and it has been highlighted as a key component of good quality guidelines internationally (1). There are several compelling reasons for engaging stakeholders.

***Evidence is imperfect.*** Clinical guidelines should be based on the best evidence. However, evidence is often of low or very low quality (9), complex to interpret, and rarely complete (10). Also, the quality of research evidence differs widely and evidence from published literature does not always encapsulate the views and experiences of the community, or address appropriate outcomes (2). Stakeholders can provide valuable evidence such as in the form of personal testimonies from patients, or views from patient organizations. For example, the Alpha-One Foundation (http://www.alphaone.org/) and the COPD Foundation (http://www.copdfoundation.org/) provide useful resources not available in the published literature.

***Recommendations are constructed through a deliberative process.*** Unlike systematic reviews that focus on assessing and analyzing evidence, guidelines make recommendations with the aim of influencing practice (11). However, evidence rarely directly translates into recommendations. Recommendations are arrived at through a deliberative process that incorporates judgments and includes consideration of the quality of evidence, the balance between desirable and undesirable effects, values and preferences, and cost (12–14). A deliberative process is participative; it involves eliciting and combining various types of evidence to reach an evidence-based judgment. In medicine, there are increasing calls for a move toward decision-making that is more inclusive and democratic, reflecting the notion that people should be involved in their own governance (10).

***The process needs to be transparent.*** Stakeholder participation is a goal in itself by encouraging participative engagement through public accountability and transparency (15). Failure to make potential conflict of interest transparent in preparing the guideline is a valid criticism (16, 17). Involving the public makes the process transparent by opening it to scrutiny, through formal consultation. The National Institute for Health and Clinical Excellence (NICE)'s guidelines are developed through an open process that takes account of the views of all those who might be affected by the guideline (usually including healthcare professionals, patients and their caregivers, service managers, hospitals, the wider public, government, and the healthcare industries) to produce credible and robust guidelines (18, 19).

***Guidelines are intended to be used.*** Guidelines are interventions that aim at improving outcomes. To achieve this, they need to be used in practice. A guideline that excludes the perspective of stakeholders may not be followed because it may fail to take into account perspectives and interests or barriers that may hamper its implementation. For example, barriers have been identified in the use of spirometry to classify severity of COPD in primary care, especially in staff training, patient acceptance, and cost and reimbursement for these tests (17). Consultation with stakeholders during guideline production can enhance uptake because it engenders a sense of ownership by addressing their concerns. The Confronting COPD survey showed that COPD is a huge burden on health care services utilization (20). Despite international authorship, many of the recommendations in the Global Initiative for Chronic Obstructive Lung Disease (GOLD) and ATS/ERS guidelines remain difficult to implement for most healthcare systems, especially about requirements for spirometry in diagnosis and staging, expensive inhaled medications and pulmonary rehabilitation, and consideration for lung volume reduction surgery (21–23). There is some evidence that input from those who deliver and manage care can enhance uptake of guidelines by addressing service delivery issues during development (24).

***Guidelines can have policy status.*** Guidelines developed by national/federal governments or even professional organizations can acquire the status of policies if they influence the behavior of institutions, funding organizations, and service organizations. This underlines the importance that clinical practice guideline (CPG) developers be accountable, not only to patients, but also to the general public (25, 26). For example, in the United States, a guideline on the management of COPD calling for pulmonary rehabilitation could ultimately be used to advocate for national coverage of this service for federal or private insurers. This would have a huge impact on whole populations and would need their consultation before being published.

***Legal considerations.*** Guidelines, if widely adopted, can have a significant impact on aspects of clinical practice. For example, they may affect the use of pharmaceutical products. If pharmaceutical companies were not consulted during creation of the guidelines or the development process was not transparent, and the pharmaceutical companies suffered from widespread adoption of the guidelines, then pharmaceutical companies may seek legal action (27). In Canada, there was a reported case of a pharmaceutical company threatening legal action against a guideline development committee after taking issue with the draft prescribing guidelines that the committee had developed (28).

Guidelines are also frequently used in the court of law and will be interpreted according to whose case needs to be made (the plaintiff or the defendant). Clinicians, patients, third-party payers, institutional review committees, other stakeholders, or the courts should never view recommendations as dictates. Guidelines or recommendations cannot take into account all of the often-compelling unique features of individual clinical circumstances that lead to certain decisions. Therefore, nobody charged with evaluating clinicians' actions should apply the recommendations in these guidelines as rote or in a blanket fashion.

### 3. At What Stage Should Stakeholders Contribute to Guidelines?

There are several points at which stakeholders can be involved:

***Selecting a topic and scoping the guideline.*** A system that allows the public to suggest topics helps ensure that guidelines address areas of importance. The Scottish Intercollegiate Guidelines Network (SIGN) and NICE encourage any group or individual to propose a guideline topic (29, 30).

Determining the frame and boundaries of the work needs input from all of those concerned, as they can influence which areas the guideline should focus on or prioritize. As an example, NICE has a system whereby stakeholders register for each clinical guideline. They include local health organizations and national organizations representing commercial, professional, service, and patient groups. Stakeholders are invited to comment on the proposed scope for a period of 4 weeks. Their comments are scrutinized by the National Collaborating Centre responsible for developing the guideline who also responds to their comments. These are posted publicly on the NICE website (31, 32).

***Contributing to the guideline development group.*** Participation of stakeholders in the guideline development group ensures that they have an equal voice at the core of the development process (6). Stakeholders may be invited to nominate expert members for the guideline group, to address specialist areas of practice. Guideline development is covered in detail in the third article of this series (3).

***Submitting evidence.*** Even if the guideline development group is thorough with its literature search, there may be evidence that is relevant to the clinical questions that has not been found. This may include ongoing research, studies published as abstracts, data on adverse effects, economic models, or studies about the experiences of patients, caregivers, or healthcare professionals. Stakeholders can submit valuable evidence for consideration by the guidelines group (19). Because of the risk of bias, this evidence needs to be assessed in the same way as the other evidence that the guideline group may have extracted through the usual systematic review process.

***Commenting on draft recommendations and peer review.*** Stakeholders contribute a specific, practice-based analysis by commenting on the validity and acceptability of the draft recommendations in a way that balances the science of guideline methodology with both how the research evidence has been interpreted (33) and the practical implications of the recommendations (34). For example, SIGN holds a public meeting with 150 to 300 stakeholders at which the recommendations are debated and stakeholders provide both feedback and suggestions for additional evidence that they might consider, or alternative interpretation of that evidence (29). The National Health and Medical Research Council (NHMRC) requires that all draft guidelines be submitted for public consultation (35). In the NICE COPD guidelines (update), the draft recommendations were emailed to registered stakeholders for comments and posted on the NICE website. Issues arising from this consultation were discussed at a further group meeting. Comments were considered, responded to, and the guideline revised (36).

***Commenting on guideline products.*** Even if stakeholders have commented on the guideline recommendations, they may not agree with the end product (such as the algorithm, quick reference guide). Inviting them to comment on the actual product can highlight problems of the presentation and content, and it can provide invaluable pointers for supporting dissemination and implementation. It may also highlight barriers in terms of resources and changes in practice, and alert the guideline developers to potential difficulties in uptake from users.

### 4. What Are the Potential Barriers to Integrating Stakeholder Involvement?

An inclusive approach needs to balance the different expectations of stakeholders, power relations, and the possibility of conflict, especially as the role of each stakeholder in guideline development is to represent and promote his/her own interests (37).

***Stakeholders have different perspectives.*** Different stakeholders value aspects of treatment quite differently (38). Patients place high value on quality of life and functional status; physicians place high value on the prevention of progression of disease and treatment compliance; and healthcare payers take a societal perspective, focusing on decisions that affect the whole population and cost implications to services. These different values impact the weights different stakeholders put on different outcomes. van der Molen argues that to improve COPD management, it is necessary to first understand the outcomes of importance to each relevant stakeholder group and, second, to refocus the measures in terms that all stakeholders can value (39). The role of the guideline development group is critical in ensuring that these diverse perspectives are taken into account in making decisions. This is covered in more detail in the eighth paper of this series (3).

***Stakeholders may be biased.*** Stakeholders have their own interests at heart and they may use those in different ways. For example, pharmaceutical companies may use the guidelines to lobby their products. Patients may have biased views informed by their own experience of the disease, which may threaten their impartiality (40). Clinical experts and professional organizations are themselves stakeholders and sometimes are clinically biased. In addition, there may be occasions on which stakeholders' attempts are made (directly or indirectly) to influence the decisions of an organization's advisory bodies through lobbying in a way that is not in the broad public interest. To balance these biases, an independent peer review system allows the academic community to comment on the robustness of the recommendations and the accuracy of interpretation of the evidence.

***The process can be costly.*** Involving stakeholders in decision-making demands commitment from the entire organization, specific managerial arrangements, and sufficient backup. This can prove to be a burden for organizations with insufficient funding (41). For example, NICE guidelines average 200 registered stakeholder organizations that comment on the scope and drafts. Administering, collating, and responding to these comments takes over 4 weeks of several people's time, including the guideline group chair.

### 5. How Can Stakeholders Be Involved Effectively?

Involving stakeholders is a challenging task that requires careful undertaking so it does not become a tokenism.

***Informing/educating stakeholders.*** Given that stakeholders have different perspectives, it is important that they are well informed about what is needed and about their roles/boundaries and input. It is also important that they understand the process to contribute effectively. For example, as part of the Consumers United for Evidence Based Health (CUE) initiative, the United States Cochrane Centre provides an online course for consumer advocates to help them understand the basic concepts of evidence-based healthcare and to help them critically appraise information found in clinical guidelines (42).

***Establishing clear communication.*** There are different ways of engaging and communicating with stakeholders. This can be done through public meetings. During the consultation on the scope of NICE guidelines, stakeholders are invited to attend a meeting at which the development process and the scope are

presented and discussed. However, stakeholders have to submit their comments in writing using the usual procedure using a standard proforma. Standard templates can help reduce unnecessary or unreasonable comments (30).

***Treating stakeholder comments equitably.*** For stakeholder involvement to succeed, it needs to be trusted by the public at large (43) and perceived to be fair so that groups do not feel disenfranchised. Commenting on draft guidelines can be a lengthy process and busy clinicians or patients groups may lack the time to scrutinize voluminous documents compared with pharmaceutical companies. Giving adequate time to respond is important for not alienating important voices from disadvantaged groups who do not have equal resources.

Likewise, it is important that stakeholder comments be treated fairly and equitably. NICE established independent Guidelines Review Panels (GRPs) to ensure the guideline developers have addressed and responded to stakeholder comments appropriately, both on the scope and on the draft guideline (18). Their role was to cross-check changes made against stakeholder comments and the developers' response to these. All of the comments and responses that were made on the draft COPD guidelines are available on the NICE website (*see* http://guidance.nice.org.uk/page.aspx?o=227392).

### 6. Should Anyone Be Excluded from the Process?

Stakeholders' input may not be appropriate because they have conflicts of interest that would bias the content of a guideline. This is particularly important when stakeholders are members of a guideline development group and have the responsibility for making the final decision. For example, a manufacturer of spirometric equipment may be a stakeholder in recommendations for office-based spirometry, but conflict of Interest issues would naturally prevent them from serving on the writing committee.

There is growing unease about the close interaction between clinical guideline authors and the pharmaceutical industry and calls for appropriate disclosure and management of financial conflicts of interest for guideline authors (42). NICE guidelines exclude pharmaceutical industries and manufacturers from its guideline development groups (GDGs). GDG conflicts of interest are recorded at the start of all guidelines, throughout their development, and are now published in the full version (18). This has an important bearing for the legitimacy and credibility of the guideline. The article by Boyd and colleagues in this issue will discuss issues related to conflicts of interest (17).

## CONCLUSIONS

Involving stakeholders is increasingly recognized as an important part of producing credible, rigorous, and transparent clinical guidelines. The Guideline International Network (GIN), an international network of 85 guideline organizations, has created the GIN Patient and Public Involvement working group reflecting the increasing recognition of this issue (26). Given the huge public health burden that COPD represents, there is a pressing need for engaging stakeholders in the development of guidelines to manage the disease effectively, especially to address outcomes that are relevant for all stakeholders. Until recently, COPD guidelines have relied largely on consensus from clinical experts and distinguished professional organizations, without sufficient involvement from stakeholders (44, 45). However, there is a shift toward better inclusiveness and transparency. There are compelling arguments for achieving effective stakeholder involvement at different stages of guideline development and consultation, but there is a paucity of evidence on what strategies work best (46). However, if performed sensitively, it can be effective (47). To work well, it needs to be thorough, transparent, fair, and inclusive. A review of the NICE guideline program by the World Health Organization (WHO) stated that "collaboration with stakeholders in the development of the guidelines through the consultation and feedback mechanisms available was in general very effective" (48). To be successful and trusted, such a process requires commitment from the entire organization, specific managerial arrangements, and adequate resources. It also relies on understanding the circumstances under which stakeholder involvement is most likely to be effective (49).

***Author Disclosures*:** F.C. consulted with and received lecture fees from NICE (National Institute for Clinical Excellence, U.K.). J.A.W. was on advisory committees of Bayer, Chiesi, GlaxoSmithKline, Novartis and Nycomed Takeda, and received lecture fees from Bayer, Boehringer Ingelheim, GlaxoSmithKline, Novartis and Nycomed Takeda. She received grants from Chiesi, GlaxoSmithKline, Novartis and Nycomed Takeda. M.K. reported no commercial interests relevant to subject matter. J.C. was employed by the American Thoracic Society. R.K. and J.W. reported no commercial interests relevant to subject matter. H.J.S. was a member of the GRADE Working Group, which received honoraria and lecture fees regarding GRADE. All workshop participants disclosed receipt of honoraria and travel reimbursement from the American Thoracic Society.

### References

1. Brouwers MC, Kho ME, Browman GP, Burgers J, Cluzeau F, Feder G, Fervers B, Graham ID, Hanna SE, Makarski J; AGREE Next Steps Consortium. Validity assessment of items and tools to support application: development steps towards the AGREE II – Part 2. *Can Med Assoc J* 2010;182:E472–E478.
2. Kelson M, Akl EA, Bastian H, Cluzeau F, Randall J, Curtis JR, Guyatt GH, Montori V, Oliver S, Schünemann HJ. Integrating values and consumer involvement in guidelines with the patient at the center. Article 8 in integrating and coordinating efforts in COPD guideline development. An official ATS/ERS workshop report. *Proc Am Thorac Soc* 2012;9:262–268.
3. Kunz R, Fretheim A, Cluzeau F, Wilt T, Qaseem A, Lelgemann M, Kelson M, Guyatt G, Schünemann HJ. Guideline group composition and group processes. Article 3 in integrating and coordinating efforts in COPD guideline development. An official ATS workshop report. *Proc Am Thorac Soc* 2012;9:229–233.
4. Vermeire P. The burden of chronic obstructive pulmonary disease. *Respir Med* 2002;96:S3–S10.
5. Mannino DM, Buist SA. Global burden of COPD: risk factors, prevalence and future trends. *Lancet* 2007;370:765–773.
6. Kelson M. The NICE Patient Involvement Unit. *EBPH* 2005;9:304–307.
7. Nilsen ES, Myrhaug HT, Johansen M, Oliver S, Oxman AD. Interventions for promoting consumer involvement in developing healthcare policy and research, clinical practice guidelines and patient information material (Cochrane Review). In: The Cochrane Library, Issue 1.Chichester, UK: John Wiley & Sons, Ltd; 2005.
8. Grol R, Dalhuijsen J, Thomas S, in 't Veld C, Rutten G, Henk Mokkink G. Attributes of clinical guidelines that influence use of guidelines in general practice: observational study. *BMJ* 1998;317:858–861.
9. Guyatt G, Akl E, Oxman A, Wilson K, Puhan MA, Wilt T, Gutterman D, Woodhead M, Antman E, Schünemann HJ. Synthesis, grading and presentation of evidence in guidelines. Article 7 in integrating and coordinating efforts in COPD guideline development. An official ATS workshop report. *Proc Am Thorac Soc* 2012;9:256–261.
10. Lomas, J, Culyer T, McCutcheon C, McAuley L, Law S; Canadian Health Services Research Foundation. Conceptualizing & combining evidence for health system guidance: final report. *CHSRF* May 2005.
11. Hayward RS, Wilson MC, Tunis SR, Bass EB, Guyatt G. Users' guides to the medical literature. VIII. How to use clinical practice guidelines. A. Are the recommendations valid? The Evidence-Based Medicine Working Group. *JAMA* 1995;274:570–574.
12. Battista RN, Hodge MJ, Vineis P. Medicine, practice and guidelines: the uneasy juncture of science and art. *J Clin Epidemiol* 1995;48:875–880.
13. Culyer AJ, Lomas LJ. Deliberative process and evidence-informed decision making in healthcare: do they work and how might we know? *Evidence and Policy* 2006;2:357–371.

14. Schünemann HJ, Woodhead M, Anzueto A, Buist S, Macnee W, Rabe KF, Heffner J. A guide for guidelines for professional societies and other developers of recommendations. Introduction to in integrating and coordinating efforts in COPD guideline development. A workshop report. *Proc Am Thorac Soc* 2012;9:215–218.
15. Stern PC, Fineberg HV, editors. Committee on Risk Characterization, National Research Council. Understanding risk: information decisions in a democratic society. Washington: National Academies Press; 1996.
16. Rabe K. Guidelines for chronic obstructive pulmonary disease treatment and issues of implementation. *Proc Am Thorac Soc* 2006;3:641–644.
17. Boyd E, Baumann M, Curtis JR, Field M, Jaeschke R, Osborne M, Schünemann H. Guideline funding and conflicts of interest. Article 4 in integrating and coordinating efforts in COPD guideline development. An official ATS workshop report. *Proc Am Thorac Soc* 2012;9: 234–242.
18. Hill J, Bullock I, Alderson P. A summary of the methods that the National Clinical Guidelines Centre uses to produce clinical guidelines for the National Institute for Health and Clinical Excellence. *Ann Intern Med* 2011;154:752–757.
19. The National Institute for Health and Clinical Excellence. How NICE clinical guidelines are developed: an overview for stakeholders, the public and the NHS, 4th edition. *NICE*, London, December 2008. Available from: http://www.nice.org.uk/aboutnice/howwework/developingniceclinicalguidelines/clinicalguidelinedevelopmentmethods/clinical_guideline_development_methods.jsp?domedia=1&mid=62F02D9B–19B9–E0B5–D4A26EC9A934FDC7
20. Rennard S, Decramer M, Calverley PMA, Pride NB, Soriano JB, Vermeire PA, Vestbo J. Impact of COPD in North America and Europe in 2000: subjects' perspective of Confronting COPD International Survey. *Eur Respir J* 2002;20:799–805.
21. Mapel D, Chen JC, George D, Halbert RJ. The cost of chronic Obstructive pulmonary disease and its effect on managed care. *Manag Care Interface* 2004;17:61–67.
22. Global Strategy for the Diagnosis, Management and Prevention of COPD, Global Initiative for Chronic Obstructive Lung Disease (GOLD). 2006. Available from: http://www.goldcopd.org.
23. Celli BR, MacNee W; Committee Members. Standards for the diagnosis and treatment of patients with COPD: a summary of the ATS/ERS position paper. *Eur Respir J* 2004;23:932–946.
24. Pierson DJ. Clinical practice guidelines for chronic obstructive pulmonary disease: a review and comparison of current resources. *Respir Care* 2006;51:277–288.
25. Eddy DM. Practice policies: where do they come from? *JAMA* 1990;263: 1265–1275.
26. Légaré F, Boivin A, van der Weijden T, Packenham C, Tapp S, Burgers J. A knowledge synthesis of patient and public involvement in clinical practice guidelines: study protocol. *Implement Sci* 2009;4:30.
27. Brushwood DB. Clinical practice guidelines and the standard of care. *Am J Health Syst Pharm* 2000;57:159–161.
28. Shuschman M. Drug company threatens legal action over Canadian guidelines. *BMJ* 1999;319:1388.
29. Scottish Intercollegiate Guidelines Network. SIGN 50: A guideline developers' handbook. Section 8: Consultation and peer review. Revised Edition January 2008 (accessed 20 June 2011). Available from: http://www.sign.ac.uk/guidelines/fulltext/50/section3.html.
30. National Institute for Health and Clinical Excellence (NICE). Topic selection (accessed 20 June 2011). Available from: http://www.nice.org.uk/getinvolved/topicselection/topicselection.jsp.
31. Wedzicha JA. Peer review and NICE COPD guidelines. *Thorax* 2004;59: 183.
32. National Collaborating Centre for Chronic Conditions. Chronic obstructive pulmonary disease: national clinical guideline on management of chronic obstructive pulmonary disease in adults in primary and secondary care (updated June 2010; accessed 20 June 2011). *Thorax* 2004; 59. Updated June 2010. Available from: http://guidance.nice.org.uk/CG101/Guidance/pdf/English
33. Fretheim A, Schünemann HJ, Oxman AD. Improving the use of research evidence in guideline development: 3. Group composition and consultation process. *Health Res Policy Syst* 2006;4:15.
34. Burgers JS, Grol RPTM, Zaat JOM, Spies TH, van der Bij AK, Mokkink HGA. Characteristics of effective clinical guidelines for general practice. *Br J Gen Pract* 2003;53:15–19.
35. National Health and Medical Research Council and Consumers' Health Forum of Australia. Statement on consumer and community participation in Health and Medical Research. Canberra: AusInfo; 2001.
36. National Institute for Health and Clinical Excellence (NICE). Chronic obstructive pulmonary disease (update): consultation table with developers responses (accessed 20 June 2011). Available from: http://www.nice.org.uk/guidance/index.jsp?action=download&o=48314.
37. Ilott, Irene; Mead, Judy; Roberts, Julia; Hammond, Ralph; Pearmain, Henny. Exploring the involvement and perceptions of professional stakeholders on NICE clinical guidelines. *Evidence & Policy: A Journal of Research, Debate and Practice* 2006;2:333–356.
38. Berberabe T, Berberabe T. Can physician and health plan get together over guidelines? *Manag Care* 1948;13:42–46.
39. van der Molen T, Pieters W, Bellamy D, Taylor R. Measuring the success of treatment for chronic obstructive pulmonary disease–patient, physician and healthcare payer perspectives. *Respir Med* 2002;96:S17–S21.
40. Boote J, Telford R, Cooper C. Consumer involvement in health research: a review and research agenda. *Health Policy* 2002;61:213–236.
41. Culyer AJ. Involving stakeholders in healthcare decisions–the experience of the National Institute for Health and Clinical Excellence (NICE) in England and Wales. *Healthc Q* 2005;8:56–60.
42. U.S. Cochrane Center. Consumers United for Evidence-based Healthcare (accessed 20 June 2011). Available from: http://us.cochrane.org/consumers-united-evidence-based-healthcare-cue.
43. Choudhry NK, Stelfox HT, Detsky AS. Relationships between authors of clinical practice guidelines and the pharmaceutical industry. *JAMA* 2002;287:612–617.
44. Nurmagambetov T, Atherly A, Williams S, Holguin F, Mannino DM, Redd SC. What is the cost to employers of direct medical care for chronic obstructive pulmonary disease. *COPD* 2006;3:203–209.
45. Smith BJ, Hender K, Frith P, Crockett AJ, Cheok F, Spedding S. Systematic assessment of clinical practice guidelines for the management of chronic obstructive pulmonary disease. *Respir Med* 2003;97:37.
46. Lacasse Y, Ferreira I, Brooks D, Newman T, Goldstein RS. Critical appraisal of clinical practice guidelines targeting chronic obstructive pulmonary disease. *Arch Intern Med* 2001;161:69–74.
47. Nilsen ES, Myrhaug HT, Johansen M, Oliver S, Oxman AD, Nilsen ES, Myrhaug HT, Johansen M, Oliver S, Oxman AD. Methods of consumer involvement in developing healthcare policy and research, clinical practice guidelines and patient information material. *Cochrane Database Syst Rev* 2010;3:CD004563.
48. World Health Organization. The clinical guideline programme of the National Institute for Health and Clinical Excellence (NICE): a review by the World Health Organization. May 2006. Copenhagen: WHO, 200 (EUR/05/5063284).
49. Boivin A, Currie K, Fervers B, Gracia J, James M, Marshall C, Sakala C, Sanger S, Strid J, Thomas V, *et al*; G-I-N PUBLIC. Patient and public involvement in clinical guidelines: international experiences and future perspectives. *Qual Saf Health Care* 2010; 19:e22.

# How to Integrate Multiple Comorbidities in Guideline Development

## Article 10 in Integrating and Coordinating Efforts in COPD Guideline Development. An Official ATS/ERS Workshop Report

**Leonardo M. Fabbri, Cynthia Boyd, Piera Boschetto, Klaus F. Rabe, A. Sonia Buist, Barbara Yawn, Bruce Leff, David M. Kent, and Holger J. Schünemann; on behalf of the ATS/ERS Ad Hoc Committee on Integrating and Coordinating Efforts in COPD Guideline Development**

***Background*: Professional societies, like many other organizations around the world, have recognized the need to use more rigorous processes to ensure that health care recommendations are informed by the best available research evidence. This is the 10th of a series of 14 articles that were prepared to advise guideline developers in respiratory and other diseases. This article deals with how multiple comorbidities (co-existing chronic conditions) may be more effectively integrated into guidelines.**

***Methods*: In this review we addressed the following topics and questions using chronic obstructive pulmonary disease (COPD) as an example. (*1*) How important are multiple comorbidities for guidelines? (*2*) How have other organizations involved in the development of guidelines for single chronic disease approached the problem of multiple comorbidities? (*3*) What are the implications of multiple comorbidities for pharmacological treatment? (*4*) What are the potential changes induced by multiple comorbidities in guidelines? (*5*) What are the implications of considering a population of older patients with multiple comorbidities in designing clinical trials? Our conclusions are based on available evidence from the published literature, experience from guideline developers, and workshop discussions. We did not attempt to examine all Clinical Practice Guidelines (CPGs) and relevant literature. Instead, we selected CPGs generated by prominent professional organizations and relevant literature published in widely read journals, which are likely to have a high impact on clinical practice.**

***Results and Conclusions*: A widening gap exists between the reality of the care of patients with multiple chronic conditions and the practical clinical recommendations driven by CPGs focused on a single disease, such as COPD. Guideline development panels should aim for multidisciplinary representation, especially when contemplating recommendations for individuals aged 65 years or older (who often have multiple comorbidities), and should evaluate the quality of evidence and the strength of recommendations targeted at this population. A priority area for research should be to assess the effect of multiple concomitant medications and assess how their combined effects are altered by genetic, physiological, disease-related, and other factors. One step that should be implemented immediately would be for existing COPD guidelines to add new sections to address the impact of multiple comorbidities on screening, diagnosis, prevention, and management recommendations. Research should focus on the possible interaction of multiple medications. Furthermore, genetic, physiological, disease-related, and other factors that may influence the directness (applicability) of the evidence for the target population in clinical practice guidelines should be examined.**

## INTRODUCTION

Professional societies, like many other organizations around the world, have recognized the need to use more rigorous processes to ensure that health care recommendations are informed by the best available research evidence. The end product of these processes are clinical practice guidelines (CPGs).

CPGs are systematically developed statements to assist practitioner and patient decisions about appropriate health care for specific clinical circumstances (1). Most CPGs, including guidelines for chronic obstructive pulmonary disease (COPD) (2, 3), collect the available evidence regarding a given disease and provide recommendations for the diagnosis, assessment of severity, and treatment of patients with that disease. However, COPD commonly exists in patients who often have multiple other chronic conditions (hereafter defined as multiple comorbidities) (4, 5), in particular heart failure (6), coronary artery disease (7), hypertension (8, 9), diabetes mellitus (10), metabolic syndrome (11, 12), cancer (13), cachexia (14), skeletal muscle abnormalities (15), depression (16), recurrent pulmonary infections (17, 18), or pulmonary hypertension (19). These multiple comorbidities may influence the clinical manifestations and natural history of COPD, and should be taken into account in the diagnosis, assessment of severity and prognosis, and management of COPD (5, 20–22).

In June 2007 the American Thoracic Society (ATS) and the European Respiratory Society (ERS) convened an international workshop of methodologists and researchers from around the world to coordinate efforts in guideline development using COPD as a model (23). Participants completed the work during the subsequent 4 years to develop a series of recommendations. This is the 10th of a series of 14 articles prepared to advise guideline developers in respiratory and other diseases. The goal of this paper is to describe how patients with multiple comorbidities should be addressed in guideline recommendations, and how issues related to patients with multiple comorbidities can be more effectively integrated in the development of guidelines.

## METHODS

The authors of this article addressed the questions listed in Table 1. We did not conduct a systematic review, but we searched PubMed and other databases of guidelines for existing systematic reviews and relevant research on the issue of guidelines,

---

This article is a section of "Integrating and Coordinating Efforts in Chronic Obstructive Pulmonary Disease (COPD) Guideline Development," an American Thoracic Society (ATS) and European Respiratory Society (ERS) Workshop Report. This official ATS/ERS Workshop Report was adopted by the ATS Board of Directors, August 2012, and by the ERS Executive Committee, February 2012.

Funding for this conference was made possible (in part) by 1R13HL 90485-01 from the National Institutes of Health. The views expressed in written conference materials or publications and by speakers and moderators do not necessarily reflect the official policies of the Department of Health and Human Services; nor does mention of trade names, commercial practices, or organizations imply endorsement by the U.S. Government.

C.B. is supported by the Paul Beeson Career Development Award Program (NIA K23 AG032910, AFAR, the John A. Hartford Foundation, the Atlantic Philanthropies, the Starr Foundation, and an anonymous donor), the Robert Wood Johnson Physician Faculty Scholars Program. C.B., B.L., and D.M.K. are funded by AHRQ R21 HS18597-01.

**Proc Am Thorac Soc Vol 9, Iss. 5, pp 274–281, Dec 15, 2012**

**DOI: 10.1513/pats.201208-063ST**
**Internet address: www.atsjournals.org**

**TABLE 1. QUESTIONS ADDRESSED REGARDING THE INTEGRATION OF COMORBIDITIES IN GUIDELINE DEVELOPMENT**

1. How important are multiple comorbidities for guidelines?
2. How have other organizations involved in the development of guidelines for single chronic disease approached the problem of multiple comorbidities?
3. What are the implications of comorbidities for pharmacological treatment?
4. What are the potential changes induced by comorbidities in guidelines?
5. What are the implications of a population of older patients with comorbidities in designing clinical trials?

including COPD guidelines, and comorbidities. We also consulted references from our own files. Finally, we reviewed guidelines on major chronic diseases from international organizations and examined whether they address the issue of comorbidities in their guidelines. Due to the limited literature, our conclusions are based on a combination of available evidence, the reported practices of organizations involved in developing guidelines, and workshop discussions.

## RESULTS

### 1. How Important Are Multiple Comorbidities for Guidelines?

Multiple comorbidities affect the epidemiology, pathophysiology, and care of COPD, all of which are critical issues usually addressed in clinical guidelines (24). The aging of the population and the decline in the age-specific death rates has led to an increase in the prevalence of multiple comorbidities at advanced ages (25–28). For example, in the United States, one third of Medicare beneficiaries in the 65- to 69-year-old age group and more than one half of those in the 85 or older group have three or more chronic medical conditions (29). Multiple comorbidities increase health care utilization (29–32), mortality (25, 26), worsening of quality of life (33), and disability (34–36).

Risk factors frequently have pleiotropic effects, which themselves have manifold consequences. For example, cigarette smoking is the major risk factor for COPD and is also an important risk factor for cardiovascular, cerebrovascular, and many other common chronic diseases, as well as several types of cancer (37–40). Comorbidities, such as heart failure, hypertension, diabetes mellitus and metabolic syndrome, coronary artery diseases, cachexia, skeletal muscle abnormalities, pulmonary infections, cancer, and pulmonary vascular disease cause variations in the clinical manifestations and natural history of COPD (5). For example, COPD complicates the diagnosis of chronic hear failure (CHF) and is thus associated with unrecognized and untreated CHF in $\geqslant 20\%$ of patients (6, 41–43) (Figure 1), and the impaired $FEV_1$ is a strong biomarker and risk factor of cardiovascular morbidity and mortality (44–46). Patients with COPD often have one or more component of the metabolic syndrome (11), and diabetes mellitus is independently associated with reduced lung function (47).

The presence of both COPD and cardiovascular disease may affect the diagnosis, severity assessment, and clinical manifestations of both conditions (48). For example, the evaluation of dyspnea or fatigue during exercise often depends on what diagnoses the patient already has. If patients have a diagnosis of cardiovascular disease, they are likely to undergo noninvasive cardiac imaging, increasing the likelihood of the diagnosis of heart failure on the basis of left ventricular dysfunction. Alternatively, when patients with stable COPD complain of dyspnea or fatigue during exercise, these symptoms may be attributed to COPD, and cardiac imaging may not be performed, potentially leaving the left ventricular dysfunction undetected (49). In addition, exacerbations of symptoms and hospitalization and mortality of patients with COPD may be caused more by comorbidities than exacerbations of COPD itself (7, 50). As in other diseases, comorbidities markedly affect the natural history of COPD. Patients with COPD mainly die of nonrespiratory diseases, specifically coronary artery, cerebrovascular diseases, and cancer (51–54). Furthermore, the presence of comorbidities such as depression and anxiety may independently affect symptoms and outcomes in COPD (55).

Thus, symptoms of COPD and comorbidities may be overlapping, treatments may interact, underlying pathophysiology may be shared, and the natural history of all conditions may be altered. As a consequence, guidelines for COPD (and other chronic conditions) should include consideration of multiple comorbidities.

### 2. How Have Other Organizations Involved in the Development of Guidelines for Single Chronic Disease Approached the Problem of Multiple Comorbidities?

Some recent guidelines for COPD acknowledge the importance of considering the role of multiple comorbidities for the diagnosis, clinical manifestations, severity assessment, prognosis, and management of COPD, but acknowledge the lack of evidence and specific guidance for clinicians to do so (56). Unfortunately, the guidelines provide few specific recommendations on how to modify care based on multiple comorbidities (2, 3, 57, 58). The same is true for some examples of recent guidelines for other common chronic illnesses, such as chronic heart failure (59), hypertension (60), and diabetes mellitus (61), which address poorly some comorbidities, including COPD, one at a time, but do not address the coexistence of multiple comorbidities at the same time. Cox and colleagues analyzed guidelines for five common chronic conditions (diabetes, heart failure, hypertension, osteoporosis, and stroke) in regard to the evidence used to support them and how they inform providers about patients of advanced old age with multiple chronic conditions (62). They evaluated 14 guidelines for age-specific recommendations, particularly for the identification or inclusion of frail older individuals, individuals older than 80 years of age, and individuals with multiple chronic conditions. They summarized their finding by stating that there is very low representation of individuals with advanced old age within guidelines and the studies upon which these guidelines are based. They, therefore, questioned the applicability of current chronic disease guidelines to older individuals.

Mutasingwa and colleagues conducted a content analysis of published Canadian guidelines for diabetes, dyslipidemia, dementia, congestive heart failure, depression, osteoporosis, hypertension, gastroesophageal reflux disease, chronic obstructive pulmonary disease, and osteoarthritis (63). They focused on the presence or absence of four key indicators of applicability of guidelines to elderly patients with multiple comorbidities (e.g., mentioning of older adults or people with comorbidities, time needed to treat to benefit in the context of life expectancy, and barriers to

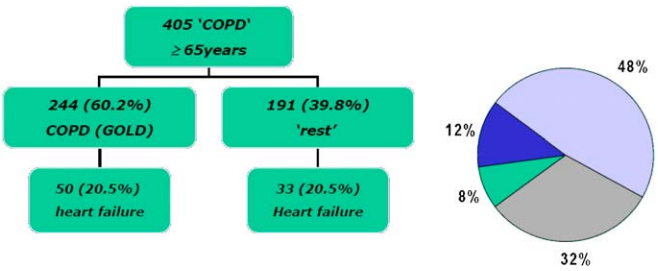


*Figure 1.* Prevalence of heart failure in stable chronic obstructive pulmonary disease (COPD) (subjects aged 65 yr or more). Data taken from Reference 49). *Pie chart*: *green*, HF only; *dark blue*, HF + COPD; *light blue*, COPD only; *gray*, negative for both HF and COPD.

implementation of the guidelines). The investigators observed that although most guidelines discuss the elderly population, few adequately address issues related to elderly patients with comorbidities (63).

There are some examples of collaborative guideline development that may serve as a model for future work to address the care of people with multiple comorbidities (23, 64, 65). The European Society of Cardiology has joined with other groups to develop recommendations for cardiovascular disease prevention in clinical practice (66). The American Geriatrics Society/California HealthCare Foundation has developed a guideline for the care of the older patient with diabetes mellitus, which extensively considers the impact of multiple comorbidities (65). The group selected six chronic conditions common in people with diabetes mellitus and reviewed guidelines and literature on each topic, developed evidence tables that summarized the data from randomized controlled trials (RCTs) on each topic, and modified existing or developed new guidelines. The panel found limited data specific to older adults with diabetes mellitus for most of the topic areas. For some areas, there were data from studies of older persons. For other areas, there were data for persons of younger ages with diabetes mellitus and the panel judged that it was reasonable to extrapolate the findings to older adults with diabetes mellitus. Recommendations were formulated as described in the two examples in Table 2. The approach chosen by the American Geriatrics Society/California HealthCare Foundation appears explicit and transparent. However, a clearer consideration for patients' values and preferences and the need for patient and clinician prioritization of the problems that should be addressed would further enhance the implementability of these guidelines as well as their relevance to everyday clinical practice. Table 3 suggests strategies for considering multiple comorbidities in the development of CPGs and patient involvement in their implementation in clinical practice. We believe that all chronic disease guidelines should have a separate section on comorbidities providing a summary of basic recommendations on diagnosis, assessment of severity, and treatment of each comorbid condition that can either be derived from other high-quality guidelines or developed *de novo*.

### 3. What Are the Implications of Multiple Comorbidities for Pharmacological Treatment?

Decisions about pharmacologic treatment represent a key area in the development of CPGs where the consideration of the impact of multiple comorbidities is crucial. A primary focus on management of a single disease may inadvertently lead to undertreatment, overtreatment, or inappropriate treatment of a patient whose health care needs may change based on the presence of multiple comorbidities (67). In particular, excess medication administration can result from adding treatments for the same condition when other causes are not considered and when there is a lack of response to therapy. This, in turn, can have unintended consequence of attempts to prevent or treat individual diseases by increasing costs, compromise adherence, and augment the risk of adverse drug events (58). Randomized clinical trials are frequently explicitly designed to exclude patients with comorbidities that may interfere with the detection of therapeutic efficacy, or which theoretically may increase the risk of adverse events (68, 69). Drugs may therefore have unanticipated effects on patients with other illnesses.

The problem of adverse side effects of medicines in patients with COPD and comorbidities is well appreciated by clinicians. For instance, systemic steroids are recommended for the treatment of exacerbations of COPD, but increase the risk of hyperglycemia in patients with COPD and diabetes mellitus (70), and may worsen osteoporosis. Conversely, β-blockers are recommended for the treatment of chronic heart failure (59, 60), but can exacerbate respiratory symptoms in patients with COPD who also have asthma (2). Bronchodilators, both β-agonists and anticholinergics, seem effective and safe in patients with COPD alone, but may increase adverse events if COPD is associated with heart failure (71) or arrhythmias.

Pharmaceutical agents can also have pleiotropic effects. Angiotensin-converting enzyme (ACE) inhibition, the cornerstone of treatment of CHF and hypertension (59, 72), may reduce mortality and morbidity in COPD (73) and improve respiratory muscle strength in patients with CHF (74). Statins, used primarily as lipid-lowering agents in the treatment of metabolic syndrome, have antiinflammatory properties that could affect co-morbidities of metabolic syndrome (e.g., COPD, CHF, and vascular diseases) (73, 75, 76).

A major reason for the lack of guidelines that address the care of people with multiple comorbidities is that the evidence on which to base the guidelines is usually very limited and indirect. RCTs are usually designed and performed for single diseases, have narrow inclusion criteria (58, 67, 69), and the populations examined frequently exclude chronic complex patients (69). More fundamentally, clinical trials are typically designed to answer a single question regarding therapeutic efficacy for a medication treating an index condition. The use of an agent with both positive and negative effects on co-existing chronic illnesses implies trade-offs that depend on the relative effects of the agent on each of the co-existing illnesses, the relative severity of the illnesses in a given patient, and patient preferences. Such questions may be difficult to answer in the context of a clinical trial. As a result, those developing clinical practice guidelines must make judgments about the degree to which the research evidence applies to patients with multiple comorbidities. Strategies can be used to account for the possible effect modification and interaction of different pharmacological agents. They can demonstrate that either the effects will differ in the population for whom the recommendation is intended from that in whom the evidence is obtained, or that there is evidence of an interaction between different interventions that would change the benefit–downside profile compared with when the interventions are administered alone. When developing recommendations for patients with COPD and multiple comorbidities, it would be ideal to evaluate the effects of the drugs in the population for whom the recommendation is intended rather than relying solely on evidence

**TABLE 2. EXAMPLE RECOMMENDATIONS FROM GUIDELINES THAT EXPLICITLY CONSIDERED MULTIPLE COMORBIDITIES**

1. "The older adult who has diabetes mellitus and hypertension should be offered pharmacological and behavioral interventions to lower blood pressure within 3 months if systolic blood pressure is 140 to 160 mm Hg or diastolic blood pressure is 90 to 100 mm Hg or within 1 month if blood pressure is greater than 160/100 mm Hg (IIIB). There are no data on the optimal timing for initiation of treatment for hypertension, but expert opinion supports the recommendation that the severity of blood pressure elevation should influence the urgency of initiating therapy. (Source guideline: 11)".
2. The older adult who has diabetes mellitus is at increased risk for major depression and should be screened for depression during the initial evaluation period (first 3 months) and if there is any unexplained decline in clinical status. (IIA)

Note: recommendations included a detailed statement about the underlying evidence that followed the recommendation. Reprinted by permission from Reference 65.

**TABLE 3. A GUIDE FOR DEVELOPMENT OF MULTIPLE COMORBIDITY CLINICAL PRACTICE GUIDELINES AND PATIENT INVOLVEMENT IN DEVELOPMENT OR APPLICATION (NOTE THAT THE EXAMPLES SHOULD NOT BE USED FOR DECISION MAKING)**

| Step | How | Example for COPD |
|---|---|---|
| Define all problems for a given patient | Ask patients (and list all problems) or review the literature on importance of problems for patients | Define which of the following is of primary concern for patients: dyspnea, depression, swelling of legs |
| Which outcome is of greatest importance to a patient with multiple co-morbidity (e.g., reducing hospitalizations, improving dyspnea) | Use tools to elicit values and preferences for that (e.g., visual analog tools, ranking exercises) | Feelsing thermometer, simple ranking techniques comparing dyspnea with fatigue and hospitalizations (described in detail) |
| Define possible options to intervene | Literature search (focus on systematic reviews), experts input on what might work | LABA, diuretics, beta-blockers, antidepressants (is the patient ready to accept few interventions only?) |
| Evaluate whether benefits or downsides (including harms) differ across populations (in particular those with different multi-morbidity) | Evaluate subgroup effects/heterogeneity across populations: use data from individual patient meta-analysis, observational studies, etc.<br>Did trials include subgroups? (use checklists of whether subgroup effects are credible).<br>Is there evidence that biology differs?<br>Make judgment about directness of the evidence | LABAs may be worse in patient with dyspnea from COPD and CHF. Treatment of dyspnea leads to improvement of depression.<br>Beta-blockers (although the evidence is not conclusive) with slightly more harm in patients with COPD and CHF |
| Evaluate greatest net benefit across populations (harms, downsides, values, and preference weighted) based on evidence profiles and present to panel making recommendations and patients | Systematically judge the expected benefits against the potential downsides after considering various interventions.<br>Explain to patients | Beta-blockers with greatest net benefit in the population of interest.<br>Treatment of depression may be of second largest net benefit. LABA and diuretic net benefit may be smaller than net benefit from beta-blockers—therefore patients having to decide for two of four medications may choose beta-blockers and antidepressants |

*Definition of abbreviations*: CHF = chronic heart failure; COPD = chronic obstructive pulmonary disease; LABA = long-acting β-agonists.

obtained from healthier patients. In the latter case, the evidence is less direct compared with evidence that directly supports recommendations, and it would influence the confidence in how the obtained effects relate to population of interest.

### 4. What Are the Potential Changes Induced by Multiple Comorbidities in Guidelines?

A critical underlying question is: How should physicians make treatment recommendations for people with multiple comorbidities, particularly if they are elderly? Realistic patient-oriented guidance requires a paradigm that incorporates these judgments (58), since clinical decision-making in such patients requires the estimation of the often subtle balance of the benefits and risks (including adverse treatment-related events) that will determine whether there are net benefits or net harms. This evaluation will frequently involve considerable uncertainty, and requires estimation of a baseline risk over a given time period. The values and preferences patients place on the treatment options and the outcomes too have to be incorporated into the decisions. These values and preferences are influenced by factors such as treatment burden and the individual's definition of quality of life. Guidelines for COPD and other diseases need to support decision making by acknowledging these factors in this complex clinical context if they are to be useful to clinicians.

The GRADE system provides a useful framework for grading both the quality of the evidence behind a recommendation and considering how strong the recommendation should be (77). Even when otherwise "high-quality" randomized studies are available, the evidence will frequently be indirect for the multimorbid population and, therefore, the quality of the evidence may be downgraded. Thus, the general effect of multiple comorbidities may be to increase the likelihood of a close or an uncertain balance between desirable and undesirable effects (risks and benefits), thus weakening the strength of the recommendations for this population.

To address these issues, comorbidities could be considered in all disease guidelines by first explicitly discussing whether patients with the most common comorbidities were included in the disease-specific trials. However, as Kravitz and colleagues have described, the determination as to whether the results of a study apply to an individual patient is not whether the patient would meet the trial inclusion criteria but whether he or she is sufficiently like, or exchangeable to, the average patient in the trial to make meaningful the resulting estimate of the average treatment effect (78). A heterogeneous sample does not eliminate concern about heterogeneity of treatment effects, because the dispersion of effects across subgroups may still be large, and analytic methods must avoid erroneous conclusions about subgroup effects (79, 80). Recommendations should be based on evidence that comes from the target population for which the guideline is intended, allowing targeting of specific recommendations to different groups within this population (58). Guidelines could be more useful if there was greater clarity in identifying exactly which of the many possible multiple morbidities were considered for which of the several recommendations within one guideline. Review of the evidence in layers considering both people with and without multiple comorbidities, as well as people at different ages, should be considered since the heterogeneity of health status regardless of the comorbidities increases with older ages. However, age alone is seldom useful in determining treatment. An older person without significant comorbid disease burden may be more likely to benefit from a therapy than a younger person with significant disease burden, or vice versa.

Second, the absolute risk reduction from a therapy for a person with one or more comorbidities must be considered, recognizing that a person with multiple comorbidities may be at either higher or lower absolute risk than the "average" person. The specific comorbidities may need to be discussed individually as the effect of the multiple comorbidities depends on the specific combinations of conditions in question. Is it known whether the relative benefit of the therapy increases or decreases in people with each combination of the multiple comorbidities? In some cases, people with multiple comorbidities may be at higher risk of a bad outcome and therefore more likely to benefit, but in other cases the risk of harm or the competing risks of dying of something else may negate or reverse the positive effects of a therapy aimed at COPD (81, 82). Thus, appropriate methods

to analyze data from heterogeneous populations are needed to understand possible variations in net treatment benefit (83).

Third, the guideline should specify the actual outcomes of each therapy, whether desired or undesired (84). If a clinician is working to apply a guideline to an individual, and is weighing and discussing the potential benefits and downsides of a therapy, it is important to have it clearly stated what the expected outcomes are (i.e., improvement in function, relief of symptoms, prevention of a stroke) (Table 3). This is not always explicit in current guidelines (58).

Fourth, the average and extremes of the length of therapy necessary to achieve this degree of risk reduction or symptom improvement should be presented. The concept of time to benefit from a therapy is essential for patients with competing risks who may have shortened life expectancy (85). The concept of "payoff time" may provide a method of tailoring guidelines to individual patients, and this will be influenced by individuals' values and preferences (83).

Fifth, guidelines should address interactions that are common or important given the prevalence of specific comorbidities. These potential interactions between a comorbidity and drugs for COPD, or between a drug for COPD and a drug for a comorbidity, or between COPD and a drug for a comorbidity, or between nonpharmacologic therapeutic recommendations, require explication.

A critical question for a patient with COPD and one or more comorbidities is what are the patient's goals or priorities for care and treatment? All of the above questions are necessary to consider in determining priorities in an individual with COPD. There is an increasing body of evidence that clinicians do not always prioritize correctly even when there is a reasonable body of evidence to guide these complex decisions (86, 87). In practice, prioritization for an individual patient requires syntheses of evidence within or across conditions. However, another critical piece must come from the patient (Table 3).

Guidelines should describe that patient preferences should always be included in discussions of goals and the selection of management decisions and that the patient's preference should be incorporated in decisions. Guidelines should provide simple summaries of risk and benefits of therapies in language that users of guidelines can communicate with patients. Recognition that patient preferences affect treatment regimens throughout the course of the disease and long before end-of-life discussions is essential. Clinicians need to know the information that they would communicate with patients such as "this therapy reduces the risk of a hospitalization for COPD of the next year from y to z for people like you" or "this therapy made 50% of people who only had COPD (without other conditions contributing to shortness of breath like you have) feel less short of breath when they walked." For example, decision analysis of the risks and benefits of warfarin use discussed with older persons with atrial fibrillation led to poor agreement with recommendations derived from guidelines, suggesting that even with excellent information and collaborative decision-making, patients may not always choose to follow guideline recommendations (88). There is often little information in guidelines on how to discuss risks, benefits in patient-friendly language to elicit preferences (89).

Feasibility, which is primarily driven by available resources, of implementing guideline recommendations must also be considered closely in the context of patients with multiple comorbidities. One facet of feasibility is medication regimen complexity (58). Methods for simplification of COPD regimens should be presented as well as discussion of the trade-offs of simplification (i.e., once per day tiotropium is more effective but also more expensive than the ipratropium 4 times per day).

Building on this, discussion of patient preferences should include the burdens of therapies and other barriers to adherence—for example, taking diuretics may make getting out and exercising or socializing difficult. Finally, how guidelines should best address comorbidities requires further study and initiatives to address this issue are underway (90).

### 5. What Are the Implications of a Population of Older Patients with Comorbidities in Designing Clinical Trials?

The patients in clinical trials that are the foundation of our current evidence base do not adequately reflect the true population of people with any chronic disease in terms of burden of multiple comorbidities (69). Similar to trials for other chronic conditions, older patients and patients with major comorbidities are specifically excluded from most clinical trials conducted in patients with COPD (91–94). Fortunately, the number of trials with explicit age exclusions for older patients has decreased. However, the percent of older patients in trials does not yet approach the percent of the overall population who are older (69, 95, 96). While age exclusions have decreased, there is some evidence to suggest that exclusions for comorbidities have increased. For example, the number of heart failure trials excluding participants with specific comorbidities increased from 1985 to 1999, with more than half of such trials excluding people with major hepatic, renal, or hematologic comorbidities (68). Again, two recent large and long COPD trials (i.e., HEALTH TORCH and UPLIFT) excluded patients with cardiovascular comorbidity (93, 94) and, thus, developing recommendations for patients with COPD and cardiovascular disease requires careful consideration of the directness of the evidence (*see* Table 3).

Exclusion and inclusion criteria are less important than who is the "average" patient in a trial; if there are few exclusion criteria, but if few people with comorbidities are actually enrolled, the results are still of questionable relevance to patients with multiple comorbidities (78). Another critical issue is that synthesizing trial results with limited generalizability to the true population with the condition may produce inappropriate guidelines for prevalent subgroups seen in practice (97) due to heterogeneity of treatment effects, defined as the "magnitude of the variation of individual treatment effects across a population" (78). A clinical trial that includes a more heterogeneous population may also see more heterogeneity of treatment effects. Average effects are not always useful, as they can represent harm to some patients, little benefit to patients who were at low risk to begin with, and a great deal of benefit to others.

Strategies for managing and understanding heterogeneity of treatment effects have been described (79, 80, 97, 98). These

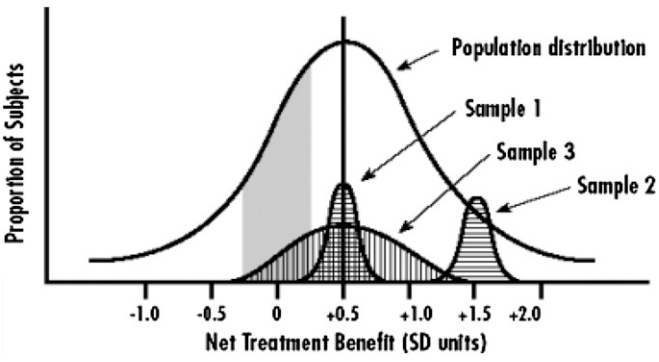


*Figure 2.* Sample 1: centered, but fails to reflect the diversity of the population. Sample 2: individuals who much more benefit from treatment than do average members of the population. Sample 3: broadly representative of the population in terms of risk, responsiveness, and vulnerability. Reprinted by permission from Reference 78.

include pretrial identification of risk groups; definition of *a priori* hypotheses; hypotheses about the direction of subgroup effects, including those at risk for poor outcomes; redesign of trials to allow for adequate power for pre-planned key subgroup analyses and analyses of heterogeneity of treatment effects; and learning from longitudinal observational studies to inform generalizability (Figure 2).

## CONCLUSIONS

Few guidelines have explicitly considered patients with multiple comorbidities (58). Detailed methods for developing recommendations for patients with multiple comorbidities are lacking. Implementing single disease guidelines presents important challenges to the clinician treating not the average clinical trial patient, but the population of patients with COPD who frequently have multiple comorbidities. We used COPD as an example for a chronic disease in this and other manuscripts in this series, and we focused mainly on nonrespiratory comorbidities. The overlap between COPD and respiratory comorbidities such as lung carcinoma, bronchiectasis, and asthma has been extensively discussed in the literature reported in COPD guidelines (54). The issues raised in this article provide a basis for a framework (Table 3) that will facilitate the integration of multiple comorbidities in the formulation and application of recommendations. We believe that it is time to tackle this issue in more depth. A critical step is the use of broader enrollment criteria and appropriate methods in randomized trials to ensure that the clinical research evidence directly addresses the populations for whom clinicians provide care in their clinical practice.

***Author Disclosures***: L.M.F. consulted with and received lecture fees from Altana, Astra Zeneca, Boehringer Ingelheim, Chiesi, GlaxoSmithKline, Merck, Novartis, Pfizer and Roche. He received grants from Altana, Astra Zeneca, Boehringer Ingelheim, Chiesi, GlaxoSmithKline, Menarini, Merck, Miat, Pfizer and UCB. C.B. reported no commercial interests relevant to subject matter. P.B. consulted with Boehringer Ingelheim. K.F.R. reported no commercial interests relevant to subject matter. A.S.B. was on advisory committees of Altana, Boehringer Ingelheim, GlaxoSmithKline, Merck, Novartis, Pfizer, Schering and Sepracor. She reported grants to the BOLD Initiative Operations Center from Astra Zeneca, Boehringer Ingelheim, Chiesi, GlaxoSmithKline, Merck, Pfizer and Schering. B.Y. consulted with Boehringer Ingelheim, GlaxoSmithKline, Novartis and Pfizer, and was on advisory committees of Astra Zeneca, Merck, Pfizer and Schering. She received grants from Astra Zeneca, Boehringer Ingelheim, Merck, Novartis and Pfizer. B.L. consulted with Amedisys Corp., the American Board of Internal Medicine and the National Quality Forum. He was a board member of the American College of Physicians and American Academy of Home Care Physicians. D.M.K. reported no commercial interests relevant to subject matter. H.J.S. was a member of the GRADE Working Group, which received honoraria and lecture fees regarding GRADE. All workshop participants disclosed receipt of honoraria and travel reimbursement from the American Thoracic Society.

### References

1. Teitelbaum D, Guenter P, Howell WH, Kochevar ME, Roth J, Seidner DL. Definition of terms, style, and conventions used in A.S.P.E.N. guidelines and standards. *Nutr Clin Pract* 2005;20:281–285.
2. Chaudhuri R, Livingston E, McMahon AD, Lafferty J, Fraser I, Spears M, McSharry CP, Thomson NC. Effects of smoking cessation on lung function and airway inflammation in smokers with asthma. *Am J Respir Crit Care Med* 2006;174:127–133.
3. Celli BR, MacNee W. Standards for the diagnosis and treatment of patients with COPD: a summary of the ATS/ERS position paper. *Eur Respir J* 2004;23:932–946.
4. Chatila WM, Thomashow BM, Minai OA, Criner GJ, Make BJ. Comorbidities in chronic obstructive pulmonary disease. *Proc Am Thorac Soc* 2008;5:549–555.
5. Fabbri LM, Luppi F, Beghe B, Rabe KF. Complex chronic comorbidities of COPD. *Eur Respir J* 2008;31:204–212.
6. Rutten FH, Moons KG, Cramer MJ, Grobbee DE, Zuithoff NP, Lammers JW, Hoes AW. Recognising heart failure in elderly patients with stable chronic obstructive pulmonary disease in primary care: cross sectional diagnostic study. *BMJ* 2005;331:1379.
7. Sidney S, Sorel M, Quesenberry CP Jr, DeLuise C, Lanes S, Eisner MD. COPD and incident cardiovascular disease hospitalizations and mortality: Kaiser Permanente Medical Care Program. *Chest* 2005;128:2068–2075.
8. Blacher J, Safar ME. Large-artery stiffness, hypertension and cardiovascular risk in older patients. *Nat Clin Pract Cardiovasc Med* 2005;2:450–455.
9. Taichman DB, Mandel J. Epidemiology of pulmonary arterial hypertension. *Clin Chest Med* 2007;28:1–22. (vii.).
10. Bolton CE, Evans M, Ionescu AA, Edwards SM, Morris RH, Dunseath G, Luzio SD, Owens DR, Shale DJ. Insulin resistance and inflammation: a further systemic complication of COPD. *COPD* 2007;4:121–126.
11. Marquis K, Maltais F, Duguay V, Bezeau AM, LeBlanc P, Jobin J, Poirier P. The metabolic syndrome in patients with chronic obstructive pulmonary disease. *J Cardiopulm Rehabil* 2005;25:226–232; discussion 233–224.
12. Poulain M, Doucet M, Drapeau V, Fournier G, Tremblay A, Poirier P, Maltais F. Metabolic and inflammatory profile in obese patients with chronic obstructive pulmonary disease. *Chron Respir Dis* 2008; 5:35–41.
13. Brody JS, Spira A. State of the art. Chronic obstructive pulmonary disease, inflammation, and lung cancer. *Proc Am Thorac Soc* 2006;3:535–537.
14. Morley JE, Thomas DR, Wilson MM. Cachexia: pathophysiology and clinical relevance. *Am J Clin Nutr* 2006;83:735–743.
15. Balasubramanian VP, Varkey B. Chronic obstructive pulmonary disease: effects beyond the lungs. *Curr Opin Pulm Med* 2006;12:106–112.
16. Hanania NA, Mullerova H, Locantore NW, Vestbo J, Watkins ML, Wouters EF, Rennard SI, Sharafkhaneh A. Determinants of depression in the ECLIPSE chronic obstructive pulmonary disease cohort. *Am J Respir Crit Care Med* 2011;183:604–611.
17. Benfield T, Lange P, Vestbo J. Chronic obstructive pulmonary disease stage and risk of hospitalization for infectious disease. *Chest* 2008;134:46–53.
18. Papi A, Bellettato CM, Braccioni F, Romagnoli M, Casolari P, Caramori G, Fabbri LM, Johnston SL. Infections and airway inflammation in chronic obstructive pulmonary disease severe exacerbations. *Am J Respir Crit Care Med* 2006;173:1114–1121.
19. Mal H. Prevalence and diagnosis of severe pulmonary hypertension in patients with chronic obstructive pulmonary disease. *Curr Opin Pulm Med* 2007;13:114–119.
20. Fabbri LM, Rabe KF. Complex chronic comorbidities: proceedings of an ers research seminar held in Rome, 11–12 February 2007 (accessed 2 July 2007). Available from: http://www.ersnet.org.
21. Brenes GA. Anxiety and chronic obstructive pulmonary disease: prevalence, impact, and treatment. *Psychosom Med* 2003;65:963–970.
22. Gift AG, McCrone SH. Depression in patients with COPD. *Heart Lung* 1993;22:289–297.
23. Schunemann HJ, Woodhead M, Anzueto A, Buist S, Macnee W, Rabe KF, Heffner J. A vision statement on guideline development for respiratory disease: the example of COPD. *Lancet* 2009;373:774–779.
24. Atkins D, Perez-Padilla R, MacNee W, Buist S, Cruz A. Priority setting in guidelines for chronic obstructive pulmonary disease: which patients and interventions need to be addressed and which recommendations need to be made? 2008.
25. Strong K, Mathers C, Leeder S, Beaglehole R. Preventing chronic diseases: how many lives can we save? *Lancet* 2005;366:1578–1582.
26. Epping-Jordan JE, Galea G, Tukuitonga C, Beaglehole R. Preventing chronic diseases: taking stepwise action. *Lancet* 2005;366:1667–1671.
27. Kirkwood TB. Understanding the odd science of aging. *Cell* 2005;120: 437–447.
28. Hadley EC, Lakatta EG, Morrison-Bogorad M, Warner HR, Hodes RJ. The future of aging therapies. *Cell* 2005;120:557–567.
29. Wolff JL, Starfield B, Anderson G. Prevalence, expenditures, and complications of multiple chronic conditions in the elderly. *Arch Intern Med* 2002;162:2269–2276.
30. Mannino DM, Braman S. The epidemiology and economics of chronic obstructive pulmonary disease. *Proc Am Thorac Soc* 2007; 4:502–506.
31. Abegunde DO, Mathers CD, Adam T, Ortegon M, Strong K. The burden and costs of chronic diseases in low-income and middle-income countries. *Lancet* 2007;370:1929–1938.
32. Gijsen R, Hoeymans N, Schellevis FG, Ruwaard D, Satariano WA, van den Bos GA. Causes and consequences of comorbidity: a review. *J Clin Epidemiol* 2001;54:661–674.

33. Bayliss EA, Ellis JL, Steiner JF. Subjective assessments of comorbidity correlate with quality of life health outcomes: initial validation of a comorbidity assessment instrument. *Health Qual Life Outcomes* 2005;3:51.
34. Fried LP, Bandeen-Roche K, Kasper JD, Guralnik JM. Association of comorbidity with disability in older women: the women's health and aging study. *J Clin Epidemiol* 1999;52:27–37.
35. Fried LP, Ferrucci L, Darer J, Williamson JD, Anderson G. Untangling the concepts of disability, frailty, and comorbidity: implications for improved targeting and care. *J Gerontol A Biol Sci Med Sci* 2004;59: 255–263.
36. Morley JE, Haren MT, Rolland Y, Kim MJ. Frailty. *Med Clin North Am* 2006;90:837–847.
37. Ezzati M, Lopez AD. Estimates of global mortality attributable to smoking in 2000. *Lancet* 2003;362:847–852.
38. Mucha L, Stephenson J, Morandi N, Dirani R. Meta-analysis of disease risk associated with smoking, by gender and intensity of smoking. *Gend Med* 2006;3:279–291.
39. Gaziano TA, Galea G, Reddy KS. Scaling up interventions for chronic disease prevention: the evidence. *Lancet* 2007;370:1939–1946.
40. Slama K. Global perspective on tobacco control: Part I. The global state of the tobacco epidemic. *Int J Tuberc Lung Dis* 2008;12:3–7.
41. Hawkins NM, Petrie MC, Jhund PS, Chalmers GW, Dunn FG, McMurray JJ. Heart failure and chronic obstructive pulmonary disease: diagnostic pitfalls and epidemiology. *Eur J Heart Fail* 2009;11:130–139.
42. Rutten FH, Cramer MJ, Lammers JW, Grobbee DE, Hoes AW. Heart failure and chronic obstructive pulmonary disease: an ignored combination? *Eur J Heart Fail* 2006;8:706–711.
43. Hawkins NM, Jhund PS, Simpson CR, Petrie MC, Macdonald MR, Dunn FG, Macintyre K, McMurray JJ. Primary care burden and treatment of patients with heart failure and chronic obstructive pulmonary disease in Scotland. *Eur J Heart Fail* 2010;12:17–24.
44. Hole DJ, Watt GC, Davey-Smith G, Hart CL, Gillis CR, Hawthorne VM. Impaired lung function and mortality risk in men and women: findings from the Renfrew and Paisley prospective population study. *BMJ* 1996;313:711–715, discussion 715–716.
45. Sin DD, Man SF. Why are patients with chronic obstructive pulmonary disease at increased risk of cardiovascular diseases? The potential role of systemic inflammation in chronic obstructive pulmonary disease. *Circulation* 2003;107:1514–1519.
46. Schunemann HJ, Dorn J, Grant BJ, Winkelstein W Jr, Trevisan M. Pulmonary function is a long-term predictor of mortality in the general population: 29-year follow-up of the Buffalo Health Study. *Chest* 2000;118:656–664.
47. Ford ES, Mannino DM. Prospective association between lung function and the incidence of diabetes: findings from the National Health and Nutrition Examination Survey Epidemiologic Follow-up Study. *Diabetes Care* 2004;27:2966–2970.
48. Boyd C, Weiss CO, Halter J, Han KC, Ershler WB, Fried LP. Framework for evaluating disease severity measures in older adults with comorbidity. *J Gerontol A Biol Sci Med Sci* 2007;62:286–295.
49. Rutten FH, Cramer MJ, Grobbee DE, Sachs AP, Kirkels JH, Lammers JW, Hoes AW. Unrecognized heart failure in elderly patients with stable chronic obstructive pulmonary disease. *Eur Heart J* 2005;26:1887–1894.
50. Anthonisen NR, Connett JE, Enright PL, Manfreda J. Hospitalizations and mortality in the lung health study. *Am J Respir Crit Care Med* 2002;166:333–339.
51. Sin DD, Anthonisen NR, Soriano JB, Agusti AG. Mortality in COPD: role of comorbidities. *Eur Respir J* 2006;28:1245–1257.
52. Mannino DM, Watt G, Hole D, Gillis C, Hart C, McConnachie A, Davey Smith G, Upton M, Hawthorne V, Sin DD, *et al*. The natural history of chronic obstructive pulmonary disease. *Eur Respir J* 2006; 27:627–643.
53. Mannino DM, Doherty DE, Sonia Buist A. Global Initiative on Obstructive Lung Disease (GOLD) classification of lung disease and mortality: findings from the Aterosclerosis Risk in Communities (ARIC) study. *Respir Med* 2006;100:115–122.
54. McGarvey LP, John M, Anderson JA, Zvarich M, Wise RA. Ascertainment of cause-specific mortality in COPD: operations of the TORCH clinical endpoint committee. *Thorax* 2007;62:411–415.
55. von Leupoldt A, Taube K, Lehmann K, Fritzsche A, Magnussen H. The impact of anxiety and depression on outcomes of pulmonary rehabilitation in patients with COPD. *Chest* 2011;140:730–736.
56. Rabe KF, Hurd S, Anzueto A, Barnes PJ, Buist SA, Calverley P, Fukuchi Y, Jenkins C, Rodriguez-Roisin R, van Weel C, *et al*. Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease: GOLD executive summary. *Am J Respir Crit Care Med* 2007;176:532–555.
57. O'Donnell DE, Aaron S, Bourbeau J, Hernandez P, Marciniuk DD, Balter M, Ford G, Gervais A, Goldstein R, Hodder R, *et al*. Canadian Thoracic Society recommendations for management of chronic obstructive pulmonary disease: 2007 update. *Can Respir J* 2007;14:5B–32B.
58. Boyd CM, Darer J, Boult C, Fried LP, Boult L, Wu AW. Clinical practice guidelines and quality of care for older patients with multiple comorbid diseases: implications for pay for performance. *JAMA* 2005; 294:716–724.
59. Swedberg K, Cleland J, Dargie H, Drexler H, Follath F, Komajda M, Tavazzi L, Smiseth OA, Gavazzi A, Haverich A, *et al*. Guidelines for the diagnosis and treatment of chronic heart failure: executive summary (update 2005): The task force for the diagnosis and treatment of chronic heart failure of the European Society of Cardiology. *Eur Heart J* 2005;26:1115–1140.
60. Mancia G, De Backer G, Dominiczak A, Cifkova R, Fagard R, Germano G, Grassi G, Heagerty AM, Kjeldsen SE, Laurent S, *et al*.; Task Force for the Management of Arterial Hypertension of the European Society of Hypertension. 2007 guidelines for the management of arterial hypertension: The task force for the management of arterial hypertension of the European Society of Hypertension (ESH) and of the European Society of Cardiology (ESC). *Eur Heart J* 2007; 28:1462–1536.
61. Ryden L, Standl E, Bartnik M, Van den Berghe G, Betteridge J, de Boer MJ, Cosentino F, Jonsson B, Laakso M, Malmberg K, *et al*. Guidelines on diabetes, pre-diabetes, and cardiovascular diseases: executive summary. The task force on diabetes and cardiovascular diseases of the European Society of Cardiology (ESC) and of the European Association for the Study of Diabetes (EASD). *Eur Heart J* 2007;28:88–136.
62. Cox L, Kloseck M, Crilly R, McWilliam C, Diachun L. Underrepresentation of individuals 80 years of age and older in chronic disease clinical practice guidelines. *Can Fam Physician* 2011;57:e263–e269.
63. Mutasingwa DR, Ge H, Upshur RE. How applicable are clinical practice guidelines to elderly patients with comorbidities? *Can Fam Physician* 2011;57:e253–e262.
64. IRCMo. International research community on multimorbidity (accessed 16 May 2008). Available from: http://www.Med.Usherbrooke.Ca/cirmo/mission_anglais.Htm.
65. Brown AF, Mangione CM, Saliba D, Sarkisian CA. Guidelines for improving the care of the older person with diabetes mellitus. *J Am Geriatr Soc* 2003;51:S265–S280.
66. Graham I, Atar D, Borch-Johnsen K, Boysen G, Burell G, Cifkova R, Dallongeville J, De Backer G, Ebrahim S, Gjelsvik B, *et al*. European guidelines on cardiovascular disease prevention in clinical practice: executive summary. Fourth Joint Task Force of the European Society of Cardiology and other societies on cardiovascular disease prevention in clinical practice (constituted by representatives of nine societies and by invited experts). *Eur J Cardiovasc Prev Rehabil* 2007; 14:E1–E40.
67. Tinetti ME, Bogardus ST Jr, Agostini JV. Potential pitfalls of disease-specific guidelines for patients with multiple conditions. *N Engl J Med* 2004;351:2870–2874.
68. Heiat A, Gross CP, Krumholz HM. Representation of the elderly, women, and minorities in heart failure clinical trials. *Arch Intern Med* 2002;162: 1682–1688.
69. Van Spall HG, Toren A, Kiss A, Fowler RA. Eligibility criteria of randomized controlled trials published in high-impact general medical journals: a systematic sampling review. *JAMA* 2007;297:1233–1240.
70. Quon BS, Gan WQ, Sin DD. Contemporary management of acute exacerbations of COPD: a systematic review and metaanalysis. *Chest* 2008;133:756–766.
71. Hawkins NM, Wang D, Petrie MC, Pfeffer MA, Swedberg K, Granger CB, Yusuf S, Solomon SD, Ostergren J, Michelson EL, *et al*. Baseline characteristics and outcomes of patients with heart failure receiving bronchodilators in the charm programme. *Eur J Heart Fail* 2010;12:557–565.
72. Mansia G, De Backer G, Dominiczak A, Cifkova R, Fagard R, Germano G, Grassi G, Heagerty AM, Kjeldsen SE, Laurent S, *et al*. 2007 ESH-ESC Guidelines for the management of arterial hypertension: the task

force for the management of arterial hypertension of the European Society of Hypertension (ESH) and of the European Society of Cardiology (ESC). *Blood Press* 2007;16:135–232.
73. Mancini GB. Clarion call for trials assessing "cardiopulmonary" agents to reduce morbidity and mortality in inflammatory lung diseases. *Chest* 2007;131:950–951.
74. Coirault C, Hagege A, Chemla D, Fratacci MD, Guerot C, Lecarpentier Y. Angiotensin-converting enzyme inhibitor therapy improves respiratory muscle strength in patients with heart failure. *Chest* 2001; 119:1755–1760.
75. Morimoto K, Janssen WJ, Fessler MB, McPhillips KA, Borges VM, Bowler RP, Xiao YQ, Kench JA, Henson PM, Vandivier RW. Lovastatin enhances clearance of apoptotic cells (efferocytosis) with implications for chronic obstructive pulmonary disease. *J Immunol* 2006;176:7657–7665.
76. van der Harst P, Voors A, van Gilst W, Bohm M, van Veldhuisen D. Statins in the treatment of chronic heart failure: biological and clinical considerations. *Cardiovasc Res* 2006;3:e333.
77. Guyatt GH, Oxman AD, Vist GE, Kunz R, Falck-Ytter Y, Alonso-Coello P, Schunemann HJ. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ* 2008;336:924–926.
78. Kravitz RL, Duan N, Braslow J. Evidence-based medicine, heterogeneity of treatment effects, and the trouble with averages. *Milbank Q* 2004;82:661–687.
79. Sun X, Briel M, Walter SD, Guyatt GH. Is a subgroup effect believable? Updating criteria to evaluate the credibility of subgroup analyses. *BMJ* 2010;340:c117.
80. Kent DM, Hayward RA. Limitations of applying summary results of clinical trials to individual patients: the need for risk stratification. *JAMA* 2007;298:1209–1212.
81. Kent DM, Alsheikh-Ali A, Hayward RA. Competing risk and heterogeneity of treatment effect in clinical trials. *Trials* 2008;9:30.
82. Varadhan R, Weiss CO, Segal JB, Wu AW, Scharfstein D, Boyd C. Evaluating health outcomes in the presence of competing risks: a review of statistical methods and clinical applications. *Med Care* 2010; 48:S96–S105.
83. Braithwaite RS, Concato J, Chang CC, Roberts MS, Justice AC. A framework for tailoring clinical guidelines to comorbidity at the point of care. *Arch Intern Med* 2007;167:2361–2365.
84. Schunemann HJ, Jaeschke R, Cook DJ, Bria WF, El-Solh AA, Ernst A, Fahy BF, Gould MK, Horan KL, Krishnan JA, *et al*. An official ATS statement: grading the quality of evidence and strength of recommendations in ATS guidelines and recommendations. *Am J Respir Crit Care Med* 2006;174:605–614.
85. Walter LC, Covinsky KE. Cancer screening in elderly patients: a framework for individualized decision making. *JAMA* 2001;285:2750–2756.
86. Hofer TP, Zemencuk JK, Hayward RA. When there is too much to do: how practicing physicians prioritize among recommended interventions. *J Gen Intern Med* 2004;19:646–653.
87. Walter LC, Bertenthal D, Lindquist K, Konety BR. PSA screening among elderly men with limited life expectancies. *JAMA* 2006;296: 2336–2342.
88. Protheroe J, Fahey T, Montgomery AA, Peters TJ. The impact of patients' preferences on the treatment of atrial fibrillation: observational study of patient based decision analysis. *BMJ* 2000;320:1380–1384.
89. Welch HG. Informed choice in cancer screening. *JAMA* 2001;285:2776–2778.
90. Boyd CM. Improving clinical practice guidelines for complex patients (accessed 15 August 2010). Available from: http://www.projectreporter.nih.gov/project_info_description.cfm?aid=7794516&icde=5163210.
91. Niewoehner DE, Rice K, Cote C, Paulson D, Cooper JA Jr, Korducki L, Cassino C, Kesten S. Prevention of exacerbations of chronic obstructive pulmonary disease with tiotropium, a once-daily inhaled anticholinergic bronchodilator: a randomized trial. *Ann Intern Med* 2005;143:317–326.
92. Aaron SD, Vandemheen KL, Fergusson D, Maltais F, Bourbeau J, Goldstein R, Balter M, O'Donnell D, McIvor A, Sharma S, *et al*. Tiotropium in combination with placebo, salmeterol, or fluticasone-salmeterol for treatment of chronic obstructive pulmonary disease: a randomized trial. *Ann Intern Med* 2007;146:545–555.
93. Wedzicha JA, Calverley PM, Seemungal TA, Hagan G, Ansari Z, Stockley RA. The prevention of chronic obstructive pulmonary disease exacerbations by salmeterol/fluticasone propionate or tiotropium bromide. *Am J Respir Crit Care Med* 2008;177:19–26.
94. Cheng JW, Nayar M. A review of heart failure management in the elderly population. *Am J Geriatr Pharmacother* 2009;7:233–249.
95. Lee PY, Alexander KP, Hammill BG, Pasquali SK, Peterson ED. Representation of elderly persons and women in published randomized trials of acute coronary syndromes. *JAMA* 2001;286:708–713.
96. Halpin DM. Lessons from the major studies in COPD: problems and pitfalls in translating research evidence into practice. *Prim Care Respir J* 2010;19:170–179.
97. Greenfield S, Kravitz R, Duan N, Kaplan SH. Heterogeneity of treatment effects: implications for guidelines, payment, and quality assessment. *Am J Med* 2007;120:S3–S9.
98. Kent DM, Kitsios G. Against pragmatism: on efficacy, effectiveness and the real world. *Trials* 2009;10:48.

# Moving from Evidence to Developing Recommendations in Guidelines

## Article 11 in Integrating and Coordinating Efforts in COPD Guideline Development. An Official ATS/ERS Workshop Report

Holger J. Schünemann, Andy D. Oxman, Elie A. Akl, Jan L. Brozek, Victor M. Montori, John Heffner, Suzanne Hill, Mark Woodhead, Doug Campos-Outcalt, Phil Alderson, Thomas Woitalla, Milo A. Puhan, Yngve Falck-Ytter, Jean Bousquet, and Gordon Guyatt; on behalf of the ATS/ERS Ad Hoc Committee on Integrating and Coordinating Efforts in COPD Guideline Development

*Introduction*: Professional societies, like many other organizations around the world, have recognized the need to use more rigorous processes to ensure that healthcare recommendations are informed by the best available research evidence. This is the 11th of a series of 14 articles that methodologists and researchers from around the world prepared to advise guideline developers for respiratory and other diseases on how to achieve this goal. For this article, we developed five key questions and updated a review of the literature on moving from evidence to recommendations.
*Methods*: We addressed the following specific questions.

1. What is the strength of a recommendation and what determines the strength?
2. What are the implications of strong and weak recommendations for patients, clinicians, and policy makers?
3. Should guideline panels make recommendations in the face of very low-quality evidence?
4. Under which circumstances should guideline panels make research recommendations?
5. How should recommendations be formulated and presented?

We searched PubMed and other databases of methodological studies for existing systematic reviews and relevant methodological research. We did not conduct systematic reviews ourselves. Our conclusions are based on available evidence, consideration of what guideline developers are doing, and pre- and postworkshop discussions.
*Results and Discussion*: The strength of a recommendation reflects the extent to which guideline developers can, across the range of patients for whom the recommendations are intended, be confident that the desirable effects of following the recommendation outweigh the undesirable effects. Four factors influence the strength of a recommendation: the quality of evidence supporting the recommendation, the balance between desirable and undesirable effects, the uncertainty or variability of patient values and preferences, and costs. Strong and weak (also called "conditional") recommendations have distinct implications for patients, clinicians, and policy makers. Adherence to strong recommendations or, in the case of weak (conditional) recommendations, documentation of discussion or shared decision making with a patient, might be used as quality measures or performance indicators.

Clinicians desire guidance regardless of the quality of the underlying evidence. Very low-quality evidence should ideally result in either appropriately labeled recommendations (i.e., as based on very low-quality evidence) or a statement that the guideline panel did not reach consensus on the recommendation due to the lack of confidence in the effect estimates. However, guideline panels often have more resources, time, and information than practicing clinicians. Therefore, they may be in a position to use their best judgments to make recommendations even when there is very low-quality evidence, although some guideline developers disagree with this approach and prefer a general approach of not making recommendations in the face of very low-quality evidence.

Guideline panels should consider making research recommendations when there is important uncertainty about the desirable and undesirable effects of an intervention, further research could reduce that uncertainty, and the potential benefits and savings of reducing the uncertainty outweigh the potential harms of not making the research recommendation. Recommendations for additional research should be as precise and specific as possible.

## INTRODUCTION

Professional societies, like many other organizations around the world, have recognized the need to use more rigorous processes to ensure that healthcare recommendations are informed by the best available research evidence. A topic of increasing interest and consensus relates to the factors beyond the quality of evidence that guideline developers should consider while developing and determining the strength of a recommendation in practice guidelines.

Guideline developers make recommendations to administer, or not administer, an intervention or management strategy on the basis of trade-offs between desirable and undesirable (i.e., including harms of an intervention and required resources) effects. If the desirable consequences outweigh the undesirable consequences, guideline panels will recommend that clinicians offer an intervention to appropriately selected patients. Conversely, if the downsides outweigh the benefits, the guidelines will recommend against the use of the intervention.

## WHAT ARE PROFESSIONAL SOCIETIES AND OTHER ORGANIZATIONS DOING NOW?

Increasingly, organizations that develop guidelines use a grading system to express the quality of evidence and the strength of a recommendation. Although professional societies use a variety of systems and methods for moving from evidence to recommendations, many of them are based on two prominent grading approaches: the system derived from the Canadian Task Force on the Periodic Health Examination (1, 2) and a successor of that system, the

This article is a section of "Integrating and Coordinating Efforts in Chronic Obstructive Pulmonary Disease (COPD) Guideline Development," an American Thoracic Society (ATS) and European Respiratory Society (ERS) Workshop Report. This official ATS/ERS Workshop Report was adopted by the ATS Board of Directors, August 2012, and by the ERS Executive Committee, February 2012.

Funding for this conference was made possible (in part) by 1R13HL 90485-01 from the National Institutes of Health. The views expressed in written conference materials or publications and by speakers and moderators do not necessarily reflect the official policies of the Department of Health and Human Services; nor does mention of trade names, commercial practices, or organizations imply endorsement by the U.S. Government.

**Proc Am Thorac Soc Vol 9, Iss. 5, pp 282–292, Dec 15, 2012**

**DOI: 10.1513/pats.201208-064ST**
**Internet address: www.atsjournals.org**

approach suggested by the Oxford Centre for Evidence-based Medicine (3). The major limitations of these systems are that they do not separate the assessment and judgments about quality of evidence from the strength of a recommendation and they rely on study design and/or risk of bias as the main indicators of quality.

Evidence alone should not determine the strength of recommendations. Guideline groups have fallen into this trap. They apply systems that do not explicitly describe the strength of a recommendation, or it is implied that the quality-of-evidence grade conveys all of the certainty or uncertainty that a guideline panel has about a given recommendation.

For instance, guidelines make statements such as "Levels of evidence are assigned to management recommendations where appropriate …. Evidence levels are indicated in boldface type enclosed in parentheses after the relevant statement – e.g., Evidence A." This guidance lacks clarity because of other factors beyond the quality of research evidence that influence the strength of recommendations; in particular, the closeness of the trade-off between the benefits and downsides of the intervention and the values and preferences of patients. When there is high-quality evidence showing a close balance of benefits and downsides, then a recommendation or its implementation will largely depend on the values and preferences people place on the management options being compared. If, however, the highest grade of the quality of evidence is assigned to such a recommendation as a sole factor for that recommendation, judgments about trade-offs and values will not be transparent, and this approach to grading recommendations may be misleading. Moreover, this approach may be misleading in relationship to the quality of the evidence if there is, for example, high-quality evidence for short-term benefits and low-quality evidence for important adverse effects and long-term effects unless there is a clear indication that the evidence grade refers to only one of the outcomes.

Consider, for instance, the use of lung volume reduction surgery (LVRS) for severe emphysema. Results of the only large-scale, well-executed randomized controlled trial to date indicate that lung resection, when combined with medical therapy, does not affect overall survival, although exercise capacity, quality of life, and other functional outcomes at 2 years were improved compared with medical therapy alone in some patient groups (4). However, surgery increases the risk of short-term mortality (5.2 vs. 1.5% at 90 d). In addition, the beneficial effects of surgery on functional outcomes appear to diminish with time. Thus, whereas some patients would be enthusiastic about undergoing LVRS because of the anticipated benefit in exercise capacity and quality of life, others who fear the risk of higher mortality in the early postsurgical phase may be less so. The Global Initiative for Chronic Obstructive Lung Disease (GOLD) states "Although the results of the large multicenter study showed some very positive results in a select group of patients, LVRS is an expensive palliative surgical procedure and can be recommended in only carefully selected patients" (5). Although most statements and recommendations in the current GOLD guidelines received a grade for the quality of evidence, this recommendation was not graded. Following the grading system used in GOLD and most existing systems, the underlying evidence would be labeled as high because it is based on randomized trials. However, fully informed patients who are offered LVRS for severe emphysema are likely to make different choices regarding this procedure; guideline panels should, despite the high-quality evidence, suggest a weak (also known as "conditional") recommendation.

In grading systems in which the grade of the recommendation depends only on the quality of evidence, a grade could be misleading if it is not accompanied by a clear description that the benefits and downsides are finely balanced. When grades are lacking, consumers of the recommendation are left with uncertainty as to what the lack of a grade indicates. Ungraded recommendations, in situations where there is uncertainty about whether the desirable consequences of adhering to the recommendation outweigh the undesirable ones, may be interpreted as strong recommendations. In other words, ungraded recommendations may be interpreted as "must not dos," or as "must dos," implemented by clinicians, and used as performance measures, even though large proportions of patients would not favor these options if they were fully informed. To overcome such issues, the American Thoracic Society (ATS), European Respiratory Society (ERS), and over 60 organizations worldwide have adopted the Grading of Recommendations, Assessment, Development, and Evaluation (GRADE) approach for evaluating the quality of evidence and grading the strength of recommendations.

In June 2007, the ATS and ERS convened an international workshop of methodologists, clinicians, and researchers from around the world to address open questions in guideline development and to coordinate efforts in guideline development for COPD and other diseases (6). Participants completed the work during the subsequent 4 years to develop advice. This is the 11th of a series of 14 articles resulting from this workshop. This article advises guideline developers for respiratory and other diseases on the factors beyond the quality of evidence that guideline developers should consider when making recommendations and determining the strength of recommendations.

Related questions about grading the quality of evidence, determining which outcomes are important, and reporting guidelines are addressed in other articles in this series (7–9).

## METHODS

The authors of this article developed and discussed the key questions described in Table 1.

We updated prior reviews of the literature on grading recommendations and factors influencing recommendations that addressed some of the key questions (10, 11). We conducted the initial update of the literature review in June 2007 and then repeated it in April 2011, searching PubMed for existing systematic reviews and relevant methodological research related to the identified questions. Search terms used for the update included "Practice Guideline"[Publication Type] OR "Guideline"[Publication Type] OR "Guidelines as Topic"[MeSH] AND evidence AND recommendations, which yielded 2,470 citations. We did not conduct systematic reviews ourselves. We also relied on prior evaluations of electronic databases and systematic reviews that suggested that the GRADE approach incorporates desired features of a grading system (11, 12).

We focused on the GRADE approach to grading the strength of recommendations, as well as the factors influencing it, because it provides guidance for going from the evidence to recommendations. And, in the context of the description of the GRADE approach, we also refer to the U.S. Preventive Services Task Force (USPSTF) approach and other approaches if they present

**TABLE 1. QUESTIONS ADDRESSED REGARDING MOVING FROM EVIDENCE TO RECOMMENDATIONS**

1. What is the "strength of a recommendation" and what determines the strength?
2. What are the implications of strong and weak recommendations for patients, clinicians and policy makers?
3. Should guideline panels make recommendations in the face of very low-quality evidence?
4. Under which circumstances should guideline panels make research recommendations?
5. How should recommendations be formulated and presented?

substantial additional factors that GRADE does not consider or if they are particularly new.

Our conclusions are based on the available evidence related to grading systems, consideration of what guideline developers are doing, workshop discussions, and the work of the GRADE Working Group (11, 13). The latter includes approximately 40 working group meetings and correspondence over the past 11 years, discussions with participants at international GRADE workshops, and feedback from users of the GRADE system (14–19). Another rationale for describing the GRADE approach in detail is that some of what has been written about GRADE is already out of date or based on an incomplete or inaccurate description of the GRADE Working Group's work (20).

## RESULTS AND DISCUSSION

### 1. What Is the Strength of a Recommendation and What Determines the Strength?

Based on the GRADE approach, the strength of a recommendation reflects the extent to which guideline developers are confident that the desirable effects of adherence to the recommendation outweigh the undesirable effects. Alternatively, if considering two or more possible management strategies (including diagnostic tests), a recommendation's strength represents the guideline developers' confidence that the net benefit clearly favors the recommended option. Desirable effects can include decreased mortality and morbidity, improved quality of life, less burden, and savings. Undesirable effects can include harms, more burden, and costs. Burdens are the demands of adhering to a recommendation that patients or caregivers (e.g., family) may dislike, such as having to adhere to a medication schedule, the inconvenience of going to the doctor's office, or having to carry around an oxygen tank to receive supplemental oxygen for respiratory disease.

Given the continuous nature of the balance of desirable and undesirable consequences, the strength of recommendation could be conceptualized on a continuous scale. Indeed, previous grading systems have sometimes used complex systems of recommendations with up to nine categories of strength of recommendations (21). GRADE has taken an approach with two categories ("strong" and "conditional/weak," where weak and conditional are alternative terms for the same category), which can be either for or against a management strategy (a total of four categories). Although in this article, we will use the terms strong and conditional (weak) recommendations, guideline panels may choose different terminology to characterize the two categories of strength, such as the term "weak" instead of the term "conditional."

When using GRADE, panel members make a strong recommendation when they are confident that the desirable effects of adherence to a recommendation outweigh the undesirable effects. Such confidence usually requires high-quality evidence providing precise estimates of benefits and downsides, and a clear difference between the magnitude of benefits and downsides (e.g., a recommendation for oxygen therapy in patients with severe COPD complicated by hypoxemia). Panel members make a weak recommendation when they believe that the desirable effects of adherence to a recommendation likely outweigh the undesirable effects, but they are not confident. Thus, if guideline developers believe that benefits and downsides are finely balanced, or appreciable uncertainty exists about the magnitude of the benefits and/or downsides, they can offer a weak recommendation for or against a management strategy (e.g., acetylcysteine for the subgroup of COPD patients not receiving inhaled steroids).

***The basis for recommendations.*** Clinical practice guidelines are intended for typical patients, but clinicians are becoming increasingly aware of the importance of individual patient values and preferences and other factors in clinical decision making. In other words, although different guideline developers who attempt to achieve complete and unbiased summaries face the same evidence, factors that influence the strength of recommendation may be much more context and patient dependent. Approaches to integrating patient values and preferences have limitations, and appropriate methods should be applied (e.g., decision aids on the individual level or population-based values on a population level). However, the basis for an evidence-based guideline is a complete summary of the evidence, ideally presented as an evidence profile that provides estimates of the magnitude of desirable and undesirable consequences of an intervention and our confidence in those estimates (Table 2). These evidence profiles typically would require little adjustment for guideline development by different organizations and could be prepared centrally through coordinated efforts, such as by authors of systematic reviews (6, 22, 23).

Once an adequate summary of the evidence is available, consideration must be given regarding how best to present that information to the guideline panel.

***Factors that influence the strength of a recommendation.*** QUALITY OF EVIDENCE. The first factor that determines the strength of a recommendation is the quality of the evidence (Table 3). The quality of evidence reflects our confidence in the estimates of effects. If guideline panels are uncertain of the magnitude of the benefits and harms of an intervention, it is unlikely they can make a strong recommendation for that intervention. Thus, even when there is an apparent large gradient in the balance of advantages and disadvantages, guideline developers will be appropriately reluctant to offer a strong recommendation for an intervention if the quality of the evidence is low.

For example, a systematic review found that oral bacterial extracts reduced COPD exacerbations with a relative risk of 0.66 (24). This estimate was, however, imprecise with a large confidence interval that included 1 (0.41–1.08). The reviewers also reported heterogeneity among the three included studies. The three included studies also suffered from a number of methodological limitations: no reporting of random allocation or allocation concealment in one trial and no reporting of intention-to-treat analysis in two. The imprecision, the inconsistency, and the study limitations reduce the extent to which we can be confident in the estimates of reduction of COPD exacerbation. As a consequence, we cannot be confident that the desirable effects of adherence to a recommendation to use oral bacterial extracts would outweigh the undesirable effects.

THE BALANCE BETWEEN DESIRABLE AND UNDESIRABLE EFFECTS. The second factor determining the strength of a recommendation is the balance between the desirable and undesirable consequences of the alternative management strategies, based on the best estimates of those consequences. Consider inhaled corticosteroids as a treatment option for a 65-year-old patient with mild COPD and frequent exacerbations. This individual's risk for suffering an exacerbation in the next year may be 20%. Considering the relative risk reduction of inhaled corticosteroids for reducing exacerbations (relative risk, 0.76; 95% confidence interval, 0.72–0.80) and this baseline risk, one can derive a simplified absolute magnitude of the effect (25). Inhaled corticosteroids, relative to placebo, would reduce the absolute risk by approximately 4.8% [$20\% - (0.76 \times 20\%)$]. Some patients who are very averse to experiencing an exacerbation may consider the downsides of inhaled corticosteroids (oral thrush, fracture, and burden of using inhalers) well worth it. Given the relatively narrow confidence interval around the relative risk reduction, one could make a strong recommendation for using inhaled corticosteroids if all patients were equally averse to exacerbations. On the other hand, if the baseline risk for an exacerbation is 5%, the absolute risk reduction is only 1.2% [$5\% - (0.76 \times 5\%)$], but

**TABLE 2. GRADE EVIDENCE PROFILE: SHOULD RESPIRATORY REHABILITATION BE USED IN PATIENTS WITH RECENT COPD EXACERBATION?**

**Author(s):** Yngve Falck-Ytter, Jan Brozek, Milo Puhan, and Holger Schünemann.

**Date:** August 16, 2011.

**Question:** Should pulmonary rehabilitation vs. usual community care be used for COPD with recent exacerbation?

**Settings:** Outpatient.

**Bibliography:** Puhan M, Scharplatz M, Troosters T, Walters EH, Steurer J. Pulmonary rehabilitation following exacerbations of chronic obstructive pulmonary disease. *Cochrane Database Syst Rev* 2011;(10):CD005305.

| Quality Assessment | | | | | | | No. of Patients | | Effect | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| No. of studies | Design | Limitations | Inconsistency | Indirectness | Imprecision | Other Considerations | Pulmonary Rehabilitation | Usual Community Care | Relative OR (95% CI) | Absolute (95% CI) | Quality | Importance |
| Hospital admission (follow-up 3–18 mo) | | | | | | | | | | | | |
| 6 | Randomized trials | No serious limitations | No serious inconsistency | No serious indirectness | No serious imprecision | None | 20/124 (16.1%) | 51/126 (40.5%) | 0.22 (0.08–0.58) | 275 (122–353) fewer per 1,000 | ⊕⊕⊕⊕ High | Critical |
| Mortality (follow-up 3–48 mo) | | | | | | | | | | | | |
| 3 | Randomized trials | No serious limitations | No serious inconsistency | No serious indirectness | Serious* | None | 8/58 (13.8%)† | 10%† | 0.28 (0.1–0.84) | 70 (15–89) fewer per 1,000 | ⊕⊕⊕O Moderate | Critical |
| | | | | | | | | 50%† | | 281 (43–409) fewer per 1,000 | | |
| Quality of life (CRQ) dyspnea (follow-up 12 and 76 wk; measured with: CRQ‡; range of scores: 1–7; better indicated by higher values) | | | | | | | | | | | | |
| 5 | Randomized trials | No serious limitations | No serious inconsistency | No serious indirectness | Serious§ | None | 128 | 130 | — | MD 0.97 (0.35–1.58) higher | ⊕⊕⊕O Moderate | Critical |
| Quality of life (SGRQ) total (follow-up 12 and 26 wk; measured with: SGRQ‖; range of scores: 0–100; better indicated by lower values) | | | | | | | | | | | | |
| 3 | Randomized trials | No serious limitations | No serious inconsistency | No serious indirectness | Serious§ | None | 63 | 64 | — | MD 9.88 (5.37–14.4) lower | ⊕⊕⊕O Moderate | Critical |
| Ambulation (as measured by 6MWD) (follow-up 1–208 wk¶; measured with: distance in meters**; better indicated by higher values) | | | | | | | | | | | | |
| 6 | Randomized trials | No serious limitations | No serious inconsistency†† | No serious indirectness | Serious§ | None | 165 | 134 | — | MD 77.7 (12.21–143.2) higher | ⊕⊕⊕O Moderate | Critical |
| Resource use—not reported | | | | | | | | | | | | |
| 0 | — | — | — | — | — | None | — | — | — | — | | |

*Definition of abbreviations*: 6MWD = 6-min walking distance; CRQ = Chronic Respiratory Questionnaire; MD = mean difference; MID = minimal important difference; No. = number; SGRQ = St. George's Respiratory Questionnaire.

* Only 23 events total.

† Median of baseline risk in studies used for low-risk estimate. High-risk estimate of 50% based on study with 4 years of follow-up. Control group risk is based on the mean across trials.

‡ CRQ: MID 0.5, moderate effect 1.0, large effect 1.5.

§ Sample size substantially lower than 400.

‖ SGRQ: MID 4, moderate effect 8, large effect 12.

¶ 76 wk, 11 d, 6 wk, and 208 wk, in the four trials, respectively.

** 6MWD: MID is 35 (30–42) m or 10% change of baseline 6MWD.

†† Although I square is 89%, this significant heterogeneity is likely due to large differences in baseline severity (baseline 6MWD).

the possible harms and treatment burden remain unchanged. Fewer patients with lower baseline risk would make the choice of taking inhaled steroids. The right choice under such circumstances is not self-evident and is likely to differ between patients.

When advantages and disadvantages are closely balanced, a weak recommendation becomes appropriate. Furthermore, if the guideline panel judges that the balance between desirable and undesirable effects varies by baseline risk, it can issue separate recommendations for groups with different baseline risks when tools for risk stratification are available for the guideline users (26, 27).

As with all other aspects of a grading system, there exists a tension between the important goal of simplicity and the danger of oversimplification. We have presented the trade-off between advantages and disadvantages as a dichotomy, although it is a continuum. The arguments for this simplification are that it can help to make the underlying judgments more transparent, and more importantly, there are clear interpretations and implications for each category, as we will describe below.

UNCERTAINTY OR VARIABILITY OF PATIENT VALUES AND PREFERENCES. The third determinant of the strength of recommendation is uncertainty, or variability, concerning values and preferences. Given that there will always be advantages and disadvantages of alternative management strategies, how a guideline panel values benefits, risks, and burden is key to any recommendation, and the strength of the recommendation. Ideally guideline panels (and

**TABLE 3. DETERMINANTS OF STRENGTH OF RECOMMENDATION**

| Factors that can strengthen the strength of a recommendation | Comment |
|---|---|
| Quality of the evidence | The higher the quality of evidence, the more likely is a strong recommendation. |
| Balance between desirable and undesirable effects | The larger the difference between the desirable and undesirable consequences, the more likely a strong recommendation is warranted. The smaller the net benefit and the lower certainty for that benefit, the more likely is a weak recommendation warranted. |
| Values and preferences | The greater the variability in values and preferences, or uncertainty in values and preferences, the more likely is a weak recommendation warranted. |
| Costs (resource allocation) | The higher the costs of an intervention—that is, the more resources consumed—the less likely is a strong recommendation warranted |

**TABLE 4. EVIDENCE TO RECOMMENDATION TABLE (HYPOTHETICAL EXAMPLE): ENHANCING TRANSPARENCY WHEN MOVING FROM EVIDENCE TO RECOMMENDATIONS**

**Question/Recommendation:** Should pulmonary rehabilitation vs. usual community care be used for COPD with recent exacerbation?

**Population:** Patients with COPD and recent exacerbation of their disease

**Intervention:** Pulmonary rehabilitation versus no rehabilitation

**Setting (if relevant):** Outpatient

| Decision Domain | Judgment | | Summary of Reason for Judgment | Explanation | Subdomains Influencing Judgment |
|---|---|---|---|---|---|
| | Yes | No | | | |
| **QoE:**<br>*Is there high- or moderate-quality evidence?*<br>The higher the quality of evidence, the more likely is a strong recommendation. | X | _ | ⊕⊕⊕O (Moderate) | There is moderate- (mortality, function, and quality-of-life outcomes) to high- (hospitalizations) quality evidence. | **QoE for benefits:**<br>Moderate to high<br>**QoE for harms:**<br>Harms not explicitly evaluated, but mortality included<br>**QoE for resource use:**<br>Resource use not explicitly evaluated<br>**Key reasons for down- or upgrading?**<br>Imprecision was a reason for downgrading for most critical outcomes<br>**All critical outcomes measured?**<br>Harms and resources not explicitly evaluated |
| **Balance of benefits versus harms and burdens:**<br>*Are you confident that the benefits outweigh the harms and burden or vice versa?*<br>The larger the difference between the benefits and harms and the certainty around that difference, the more likely is a strong recommendation. The smaller the net benefit or net harm and the lower the certainty for that net effect, the more likely is a conditional/weak recommendation. | X | _ | There is considerable benefit while little clinical harm or downsides are expected. | There is a significant reduction in hospital admissions (OR, 0.22, 95% CI, 0.08–0.58) with 275 (95% CI, 122–353) fewer per 1,000 patients for a baseline risk of approximately 40%.<br>Mortality during follow-up of 3–48 mo) was significantly reduced (OR, 0.28, 95% CI, 0.1–0.84) with 70 (95% CI, 15–89) fewer per 1,000 for a control group risk of 13%.<br>Quality of life (CRQ) dyspnea, ambulation (as measured by 6-min walking distance) improved on average more in the pulmonary rehabilitation group than in the control group, and this difference exceeded minimal important difference for each of these outcomes. | **Baseline risk for benefits:**<br>• Is the baseline risk similar across subgroups?<br>• Should there be separate recommendations for subgroups?<br>**Baseline risk for harm and burden?**<br>• Is the baseline risk similar across subgroups?<br>• Should there be separate recommendations for subgroups?<br>**Relative risk for benefits and harms:**<br>• Are the relative benefits large?<br>• Are the relative harms large?<br>**Requirement for modeling:**<br>Is there a lot of extrapolation and modeling required for these outcomes? |
| **Values and preferences:**<br>*Are you confident about the assumed or identified relative values, and are they similar across the target population?*<br>The more certainty or similarity in values and preferences, the more likely a strong recommendation. | X | _ | Benefits much higher valued than expected minor harms. | A high value was placed on avoiding hospitalizations and mortality as well as improving quality of life.<br>A low value was placed on possible adverse events. | **Perspective taken:**<br>Patients or public<br>**Source of values:**<br>Guideline panel assessment<br>**Source of variability if any:**<br>Not a lot of variability<br>**Method for determining values satisfactory for this recommendation:**<br>Yes, given the expected small variability and difference between guideline panel and patients |

(*Continued*)

**TABLE 4. (CONTINUED)**

| Decision Domain | Judgment: Yes | Judgment: No | Summary of Reason for Judgment | Explanation | Subdomains Influencing Judgment |
|---|---|---|---|---|---|
| **Resource implications:**<br>*Are the resources worth the expected net benefit from following the recommendation?*<br>The lower the cost of an intervention compared with the alternative, and other costs related to the decision—that is, the fewer resources consumed—the more likely is a strong recommendation in favor of that intervention. | X | _ | Resources required are worth the net benefit considering the benefit on mortality and hospitalizations. | There are resources required to provide pulmonary rehabilitation, but these resources are worth the expected benefits, and downstream treatment costs for COPD exacerbations such as hospitalizations are avoided. | **What are the cost per resource unit?**<br>Although not evaluated here, a hospital bed per day is typically considered to be $800. Rehabilitation cost are approximately $3,000 to 5,000 per program per patient.<br>**Feasibility:**<br>Is this intervention generally available?<br>**Opportunity cost:**<br>Is this intervention and its effects worth withdrawing or not allocating resources from other interventions?<br>**Differences across settings:**<br>Is there lots of variability in resource requirements across settings? |
| **Overall strength of recommendation** | Strong | | The guideline panel recommends that patients with recent exacerbations of their COPD undergo pulmonary rehabilitation. (NOTE: this is a hypothetical recommendation developed for this article and not intended for clinical decision making.) | | |
| **Remarks** | This recommendation places a high value on the benefits that can be expected (mortality reduction, reduction in hospitalizations, and improvement in quality of life) and a relatively low value on the required resources. All patients should receive recommended usual care in addition to rehabilitation. (NOTE: this is a hypothetical recommendation developed for this article and not intended for clinical decision making.) | | | | |

*Definition of abbreviations*: CI = confidence interval; COPD = chronic obstructive pulmonary disease; CRQ = Chronic Respiratory Questionnaire; OR = odds ratio; QoE = quality of evidence.

their recommendations) would be informed by the best evidence about values and preferences that is directly obtained from patients or the public. Frequently, however, guideline panels must act on behalf of the target population and make assumptions about the underlying values and preferences. These assumptions must be transparently described. Regardless, one could argue that there is always large uncertainty about values and preferences. However, there is some systematic study of values and preferences, and clinicians' experience with patients provides additional insight. For example, consider patients with a COPD exacerbation that is severe enough to require mechanical ventilation. One could easily assume that the potential benefits of mechanical ventilation (a reduction of the rate of an almost certain mortality) outweigh the potential downsides (e.g., pneumonia and discomfort) (28). However, patients vary significantly in terms of their preferences for receiving mechanical ventilation because of the different values they ascribe to the potential benefits, downsides, or both. For this reason, it is unlikely that a guideline panel would strongly recommend mechanical ventilation for all COPD patients with severe respiratory failure, despite the clear survival benefit. A strong recommendation could be issued, however, that patients or their proxies receive information allowing an informed choice.

COSTS OR RESOURCE USE. The final determinant of the strength of a recommendation is cost or resource use. One could consider resource use as one of the outcomes when balancing positive and negative consequences of competing management strategies. Cost, however, is much more variable over time, geographic areas, and implications than are other outcomes. For example, generic drugs may be less costly than patented drugs, and charges for the same drug differ widely across jurisdictions. In addition, the implications of the used resource vary widely. For instance, a year's prescription of a drug may pay for a single nurse's salary in the United States, 10 nurses' salaries in Romania, and 30 nurses' salaries in India.

Thus, while higher costs will reduce the likelihood of a strong recommendation in favor of a particular intervention, the context of the recommendation will be critical. In considering resource allocation issues, guideline panels must thus be very specific about the setting to which a recommendation applies and the perspective that is used (e.g., which costs were considered and whether a health system or a payer perspective was taken). Furthermore, recommendations that are heavily influenced by costs are likely to change over time as resource implications change. Table 4 provides an example for transparently documenting the judgments and reasons of panels when moving from from evidence to recommendations.

***How do other organizations make recommendations?.*** Many organizations (e.g., American College of Physicians, American College of Chest Physicians, World Health Organization [WHO], and National Institutes of Health and Clinical Excellence) use the GRADE or nearly identical approaches (17, 29–32). Others, like the USPSTF approach, are very similar. For example, the USPSTF estimates the magnitude of benefits and harms, and synthesizes these assessments into an estimate of the magnitude of net benefit. It then weighs the balance of the benefits (often quantified in terms of lives extended or illness events averted) against the harms, possibly measured in terms of the health consequences of false-positive screening tests or adverse effects of treatment. The Task Force provides the example of prophylactic aspirin therapy among men, for which the benefits include fewer coronary heart events and the harms include more major gastrointestinal bleeding episodes (33). Although quantitative approaches for benefit and harm assessment exist (e.g., multicriterion decision analysis, net number of events prevented, or probabilistic simulation), the USPSTF recognizes that such analyses can be complex. Given the results of an outcomes table, the USPSTF then categorizes the magnitude of net benefit as substantial, moderate, small, or zero/negative. This last category refers to preventive measures that, if implemented in the general primary care population, can be expected to achieve no net benefit or to result in overall harm. Despite the quantity of objective evidence reviewed, the USPSTF must use judgment in determining final estimates of net benefit; outcomes tables help make this

judgment explicit and transparent. The USPSTF can rarely, if ever, assign an exact magnitude to the benefits or harms of implementing a preventive service. It can, however, put boundaries around the estimate of net benefits. The upper and lower boundary limits on the net estimated benefit make up a "conceptual confidence interval." This range is bound by the best- and worst-case scenario estimates based on available evidence. The interval is not meant to have a statistical interpretation. The GRADE approach is similar in that it considers the worst- and best-case scenario on the basis of an evidence profile and determines whether the worst-case scenario would significantly alter the recommendation; in such cases, a strong recommendation should not be offered. GRADE suggests that the assumptions should be described transparently and be based on quantitative modeling. However, so far, this is infrequently done by groups using GRADE. In the absence of modeling, the emphasis is on describing judgments in the most transparent way.

For the carotid artery stenosis screening recommendation, the USPSTF "bound" the benefits for screening a primary care population on the basis of population prevalence, screening accuracy, and treatment benefit. Randomized trials specified the maximum potential benefit from selected individuals having carotid endarterectomies performed by selected surgeons. The Task Force concluded that the magnitude of benefits in the primary care population could not be greater than the magnitude shown in the randomized controlled trials and would probably be smaller in real-world settings. In general, the Task Force believes that preventive services graded as either A or B should be provided to eligible patients, those with a C grade should not be offered routinely, and D-grade services should not be provided. Services for which the certainty of the evidence is low because of insufficient evidence about net benefit are designated using an I statement, and no recommendation is made.

The recently described Australian method for formulating and grading recommendations in evidence-based clinical guidelines (FORM) bases its recommendations, like many other systems, primarily on the trustworthiness of the underlying body of evidence based on the following five criteria: evidence base (level of studies and risk of bias), consistency, clinical impact, generalizability, and applicability (34).

Most systems other than GRADE do not provide clear criteria for when the confidence in observational studies can or should be increased, particularly in areas where randomized trials are sparse, such as public health or many of the surgical disciplines (35). GRADE, despite frequent misconceptions that it focuses on randomized trials, allows an assessment of the confidence in the estimates of effect regardless of the underlying basic study design before moving from evidence to recommendations (36). Furthermore, GRADE uniquely offers guidance for moderating confidence in the estimates in the face of problems with the overall body of evidence, such as evidence of publication bias and between-study inconsistency.

### 2. What Are the Implications of Strong and Conditional (Weak) Recommendations for Patients, Clinicians, and Policy Makers?

The advantage of GRADE's binary grading of strength of recommendations is that it provides clear direction or interpretation aids for patients, clinicians, and policy makers. The implications of a strong recommendation, such as to use (vs. not to use) antibiotics for the treatment of community-acquired pneumonia in patients with COPD, are:

- For patients and the public: Most individuals in this situation would want the recommended course of action, and only a small proportion would not. Formal decision aids are not likely to be needed to help individuals make decisions consistent with their values and preferences.
- For clinicians and other healthcare providers: Most individuals should receive the intervention.
- For policy makers: The recommendation can be adapted as policy in most situations. Adherence to this recommendation according to the guideline could be used as a quality criterion or performance indicator.

The implications of a conditional (weak) recommendation, such as to use LVRS in patients with severe (upper lobe-predominant) emphysema, are:

- For patients and the public: The majority of individuals in this situation would want the suggested course of action, but many would not. Decision aids could be useful in helping individuals make decisions consistent with their values and preferences.
- For clinicians and other healthcare providers: Different choices will be appropriate for different patients, and clinicians must help each patient arrive at a management decision consistent with her or his values and preferences. Decision aids could be useful in helping individuals make decisions consistent with their values and preferences.
- For policy makers: Policy making will require a more substantial debate and involvement of many stakeholders. Adequate documentation of the decision making process for a conditional (weak) recommendation could be used as a quality measure, in particular, if the conditional (weak) recommendation is based on high-quality evidence.

Strong recommendations may not be important from the patient's or from the health system perspective. If the choice is relatively unimportant, some patients may not bother with even strong recommendations. This is likely, for example, if they are at low risk and faced with numerous suggestions to change their lifestyle.

Policy makers and public health officials should consider a number of issues beyond the strength of a recommendation and may decide that some strong recommendations that may be important for individual patients have low priority from a system or public health perspective. These issues include the prevalence of the health problem (higher priority to more prevalent conditions); considerations of equity (higher priority to interventions that address health inequities by targeting disadvantaged populations); total cost to society (lower priority to interventions with very high total costs); and the potential for improvement in quality of care (higher priority to under- or overused interventions). Thus, if guideline panels are addressing funders or health system managers, they should make transparent the manner in which issues of prevalence, equity, cost, and improving quality of care influence their recommendations.

***Performance measures.*** Practices based on high-quality evidence for which the desirable consequences far exceed undesirable consequences with little anticipated variability in patient values and preferences constitute appropriate candidates for quality-of-care criteria. When evidence is lower quality, desirable and undesirable consequences are more closely balanced, or values and preferences vary considerably across patients, variable management is reasonable, and management practices should be considered discretionary and not candidates for quality assessment. GRADE provides guidance on these matters: The management options associated with strong recommendations are particularly good candidates for quality criteria. However, strong recommendations (and thus performance measures)

based on low- or very low-quality evidence should be rare. Most performance measures should be based on high- or moderate-quality evidence. When a recommendation is conditional (weak), discussing with patients and families the relative merits of the alternative management strategies and appropriate documentation of this interaction may become a quality criterion. The quality criterion in this case could require evidence that evidence-based information was shared with patients and that patients were able to actively deliberate and make decisions in a manner consistent with their preferences for participation.

### 3. Should Guideline Panels Make Recommendations in the Face of Very Low-Quality Evidence?

If guideline panels decide not to develop recommendations in the face of low- or very low-quality evidence, they fail one of their fundamental missions: to provide guidance and solutions for the healthcare provider who requires answers to the questions. Clinicians would be frustrated by consulting a guideline that states that there is not good enough evidence to provide a recommendation. Clinicians would be left with uncertainty despite their eagerness to obtain information about the best course of action and might quickly become discouraged. Guideline panels typically have more time, resources, and diverse types of expertise than individual clinicians. Thus, in most instances, they should use their best judgments to make specific and unambiguous recommendations (albeit conditional [weak] ones) and transparently lay out the judgments they made. Also, higher-quality evidence may never be obtained, and therefore, physicians need guidance regardless of the quality of the underlying evidence.

However, some panel members in our workshop discussion disagreed with this conclusion and believe that no recommendations should be made when the evidence is considered "insufficient." The American College of Physicians, American Academy of Family Physicians, and USPSTF all use an "insufficient evidence to make a recommendation" category. Workshop participants argued that it is too risky for a guideline panel to make a recommendation based on low- or very low-quality evidence when there is a substantial risk that the panel decision may be wrong. Other panel members felt that if this uncertainty and the potential for being wrong is made clear and transparently communicated in a conditional (weak) recommendation based on low- or very low-quality evidence, the benefits would be greater than the risk of leaving it up to individual clinicians to make the difficult decisions unassisted by advice from the panel. Nonetheless, there was agreement that an option not to make a recommendation should be available for guideline panels.

There is limited evidence about whether guideline consumers prefer knowing about the underlying strength of a recommendation and its determinants. We are aware of three such studies, one of which has not been published. The first study, conducted by UpToDate, asked a small group of users to compare graded with nongraded recommendations and explored—in a focus group setting—reasons for their answers (UpToDate, personal communication). Users of recommendations preferred knowing about the underlying quality of evidence and strength of recommendations. The second is our own study of a small group of the general public interested in healthcare issues (37). Participants preferred to know about the uncertainty relating to outcomes of a treatment or a test, but they were slightly more interested in knowing about uncertainty relating to benefits than harms (96% vs. 90%). Participants also expressed a strong preference to be informed about the quality of evidence that supports a recommendation (mean rating of 2.5 [95% confidence interval: 2.26–2.73] on a typical 7-point Likert scale from −3 to 3 where 0 is neutral). The third is a series of two randomized controlled trials, which investigated the use of summary-of-findings tables in Cochrane Collaboration's systematic reviews (that do not make recommendations) and found that users valued and understood assessments of the quality of evidence (38, 39). The required time to obtain information, as well as understanding and accessibility of the information presented in systematic reviews, all showed favorable outcomes in these two studies.

### 4. Under Which Circumstances Should Guideline Panels Make Research Recommendations?

Consider a promising intervention associated with potential side effects and substantial cost. If the evidence available about the desirable and undesirable effects of this intervention is of very low to low quality, guideline panels considering this intervention may be uncertain as to how to proceed.

As we discussed above, one option is to not formulate a recommendation at all. However, once guideline developers decide a question is worth asking, they may consider that this is an unacceptable position that abdicates their mission to provide guidance (*see above*).

Alternatively, they could decide to formulate conditional (weak) recommendations in favor of or against using the intervention in practice without indicating whether additional research is required. This has several risks: prematurely promoting the use of the intervention, potentially diverting scarce resources, and exposing patients to harm without benefit. Furthermore, the existence of the recommendation could discourage funding agencies from investing resources in the research and patients from enrolling in the studies.

The third and preferred option is for the panel to recommend that the intervention be used in the context of research complemented by guidance for what are the best management options until further research becomes available. This option may promote the conduct of research to answer important questions, thus decreasing uncertainty about the optimal way to manage patients. Panels that are considering this approach need to consider the relative merits of recommending research versus recommending for or against the intervention. There are no well-established criteria for guiding panels to make the determination of whether research should be done. Nonetheless, the following three criteria must be met for a recommendation to use an intervention in the context of research to be sensible:

1. There must be important uncertainty about the effects of the intervention (e.g., low- or very low-quality evidence for either or both the desirable and undesirable consequences).
2. Further research must have the potential to reduce that uncertainty at a reasonable cost.
3. The potential benefits and savings of reducing the uncertainty outweigh the potential harms and costs of either using or not using the intervention based on currently available evidence (40, 41).

Conducting the necessary research should be important enough to spend research resources, but it is unlikely that a guideline panel alone is able to make this decision.

The research recommendations should be detailed regarding the specific research questions that should be addressed, particularly which patient-important outcomes should be measured, and other relevant aspects of what research is needed (42). Such patient-oriented research about the effectiveness and safety of interventions will often take the form of large trials, measuring

the effects of interventions versus standard care on patient-important outcomes (mortality, morbidity, and quality of life) and following all participants over extended periods. Panels may also wish to recommend further fundamental research (seeking further biological understanding) or implementation research (addressing how the intervention can be efficiently implemented in different settings).

If guideline panels are uninformed about the available (global) research resources and the other meritorious research questions waiting for funding, and if they lack information and skills to credibly establish research priorities, they should rarely formulate research recommendations. Alternatively, the panel may exercise restraint and limit its recommendation to asking the relevant questions with sufficient detail for researchers to identify the best methods to address these and render high-quality evidence.

In some cases, guideline panels may want to formulate strong research recommendations. This may take place when the condition leads to important deterioration in patient's quality or duration of life, there are no safe and effective treatments available, the research seems feasible both scientifically and financially, and the research is likely to substantially reduce uncertainty. Another example for the formulation of strong research recommendations could be the emergence of new and far-reaching health threats.

As we discussed above, a recommendation for using an intervention only in research contexts implies a strong recommendation against using it outside of a research context. When this is the intent of the panel, both recommendations should be made explicit. In some instances, for example, when there is emerging evidence of unexpected harm in patients using an intervention that is associated with small benefit, a panel may decide to make a conditional (weak) recommendation against or for the intervention in practice and a strong recommendation for the use of the intervention in research.

Defining the population, intervention, comparator, and outcomes (PICO) explicitly will make research recommendations more helpful. Brown and colleagues expanded PICO to add the state of the evidence and time of literature search to EPICOT (42). EPICOT stands for:

- E (Evidence): What is the current evidence?
- P (Population): Diagnosis, disease stage, comorbidity, risk factor, sex, age, ethnic group, specific inclusion or exclusion criteria, clinical setting.
- I (Intervention): Type, frequency, dose, duration, prognostic factor.
- C (Comparison): Placebo, routine care, alternative treatment/management.
- O (Outcome): Which clinical or patient-related outcomes will the researcher need to measure, improve, influence, or accomplish? Which methods of measurement should be used?
- T (Time stamp): Date of literature search or recommendation.

Additional factors that would inform research recommendations include the baseline risk of the disease, the burden of the condition, and the study type that would best suit subsequent research (42).

Since the target audience for most practice guidelines is clinicians, the recommendations for research may seem misplaced and distracting among the recommendations related to practice. If this is the case, research recommendations could be placed in an appendix or special sections in the guideline directed at researchers and research funding agencies. A similar formatting decision should affect the design of executive summaries.

### 5. How Should Recommendations Be Formulated and Presented?

A number of instruments for evaluating clinical practice guidelines can be used as checklists for formulating recommendations (43, 44). The National Guideline Clearinghouse also imposes a standard for reporting on organizations that want their guidelines included in that database (45). It includes 52 items under the following headings: scope, methodology (including rating scheme and cost analysis), recommendations, evidence supporting the recommendations, benefits/harms of implementing the recommendations, contraindications, qualifying statements, implementation of the guideline, Institute of Medicine national healthcare quality report categories, identifying information and availability, and disclaimer; in addition to indexing attributes.

Similarly, some journals have standard formats for reporting clinical practice guidelines, including structured abstracts with the following headings (46):

**Objective:** a succinct statement of the objective of the guideline, including the targeted health problem, the targeted patients and providers, and the main reason for developing recommendations concerning this problem for this population.

**Options:** principal practice options that were considered in formulating the guideline.

**Outcomes:** significant health and economic outcomes identified as potential consequences of the practice options.

**Evidence:** methods used to gather, select, and synthesize evidence, and the date of the most recent evidence obtained.

**Values:** persons and methods used to assign values (relative importance) to potential outcomes of alternative practice options.

**Benefits, Harms, and Costs:** the type and magnitude of the main benefits, harms, and costs that are expected to result from guideline implementation.

**Recommendations:** a brief and specific list of key recommendations.

**Validation:** the results of any external review, comparison with guidelines developed by other groups, or clinical testing of guideline use.

**Sponsors:** key persons or groups that developed, funded, or endorsed the guideline.

Both checklists for evaluating guidelines and for reporting guidelines include items that may be dependent on the specific setting in which a guideline is developed or used. There are a number of international organizations, including the WHO, that develop guidelines that are intended to be used in a variety of settings around the world. We did not find any published articles that addressed methods for taking into account setting-specific factors in international guidelines, although several groups are working on methods for adapting guidelines developed in one setting for use in another (47).

Shekelle and colleagues evaluated the effect of different levels of specificity of recommendations on the test-ordering behavior of clinicians using clinical vignettes (48). They found that clinicians receiving nonspecific recommendations ordered fewer indicated tests than physicians receiving specific recommendations. The authors concluded that the clarity and clinical applicability of a guideline might be important attributes that contribute to the effects of practice guidelines. We did not find any other comparisons of different ways of formulating recommendations, and it is likely that the way in which recommendations are formulated may need to be adapted to the specific characteristics of a guideline. Hussain and colleagues found that there was great inconsistency within and across guidelines in how recommendations were formulated (49), and most recommendations lacked an indication about their strength. There is a consensus that recommended actions should be stated precisely.

Few written standards for how recommendations are formulated exist. However, guideline developers should provide clinicians with clear and consistent indicators of the strength of recommendations. For strong recommendations, the GRADE Working Group has suggested adopting terminology such as, "We recommend …" or "Clinicians should … ," in addition to using symbols (or numbers) and labels (e.g., "strong recommendation") (19, 31, 37). When panels make a weak or conditional recommendation, they could use less definitive wording, such as, "We suggest …" or "Clinicians might … ." Furthermore, guideline panels should describe the population (described by the disease and other identifying factors), intervention, and comparison (as detailed as feasible) to make their recommendations as specific as possible.

However, to date there is little empirical evidence about wording of the strength of recommendations. Lomotan and colleagues explored the "level of obligation" assigned to a number of terms commonly used in clinical practice guidelines (50). They concluded that participants assigned different levels of obligation to "must," "should," and "may." We compared three wording approaches, each expressing two strengths of recommendation ("we recommend"/"we suggest"; "clinicians should"/"clinicians might"; "we recommend"/"we conditionally recommend") (51). None of the approaches was clearly superior to the others in conveying the strength of recommendations.

## CONCLUSIONS

ATS, together with many other organizations around the world, including the WHO, American College of Chest Physicians, ERS, Endocrine Society, Scottish Intercollegiate Guideline Network, Infectious Disease Society of America, and UpToDate, has decided to apply the GRADE system, a grading system that (*1*) follows a common methodology for grading the quality of evidence and strength of recommendations, (*2*) is sensible, and (*3*) is being used widely (14). Representatives from these organizations have been involved in the development of this system. Some of these organizations have developed many guidelines using this approach, and it promises to become a unifying approach for many other clinical practice guidelines. Recently, other efforts, such as the Allergic Rhinitis and its Impact on Asthma guidelines, have applied the GRADE system, further harmonizing the way guidelines are developed internationally (52).

In general, the evidence that graded recommendations have advantages over nongraded recommendations is limited, but there are strong arguments for graded recommendations. These include the clear and transparent communication of how confident users can be that the desirable consequences of adherence to a recommendation outweigh the undesirable consequences. However, whether through the use of symbols or words, there may be important differences in the subjective interpretation of the concept of "strong" or "weak/conditional" recommendations. Informal work of the GRADE Working Group has revealed that different individuals and organizations prefer alternative terms to the word "weak," such as "conditional" (53). Organizations that have adopted GRADE, like the Centers for Disease Control and Prevention Advisory Committee on Immunization Practices, have chosen such alternative terms (54).

It is often said that evidence alone does not make decisions. This is correct if evidence is restricted to evidence from research studies. However, in the absence of such "research evidence," information from or related to individual patients about the clinical state and circumstances and the patient's values and preferences can be considered a form of (very low quality) evidence. In situations in which the information from single patients is considered evidence in the context of research evidence, evidence alone can and should be used and will be appropriate for decision making.

### Further Work

Further work is needed on several of the questions asked in this review, including the evaluation of methods for assessing values and preferences and how best to word recommendations. A particularly challenging area where more work is needed is the development and evaluation of methods for taking into consideration information that varies from setting to setting when making global recommendations. A great deal of operational research regarding the use of specific terminology and factors that influence how to move from evidence to recommendation is required (e.g., DECIDE collaboration: www.decide-collaboration.eu).

***Author Disclosures***: H.J.S. was a member of the GRADE Working Group, which received honoraria and lecture fees regarding GRADE. A.D.O. was a member of the GRADE Working Group. E.A.A. was a member of the GRADE Working Group. J.L.B. was a member of the GRADE Working Group. V.M.M. was a member of the GRADE Working Group. and J.H. reported no commercial interests relevant to subject matter. S.H. was employed by the World Health Organization; she was a member of the GRADE Working Group. M.W. consulted with GlaxoSmithKline and received lecture fees from Pfizer. D.C.-O. received royalties from McGraw-Hill. P.A. was employed by NICE (National Institute for Clinical Excellence, U.K.); he was a member of the GRADE Working Group. T.W. reported no commercial interests relevant to subject matter. M.A.P. received lecture fees from GlaxoSmithKline and Grants from Astra Zeneca, Boehringer Ingelheim, and GlaxoSmithKline; he was a member of the GRADE Working Group. Y.F.-Y. was a member of the GRADE Working Group. J. B. served on the ADAPTE Collaboration and as a trustee of AGREE Research Trust and Guidelines International Network. G.G. was a member of the GRADE Working Group. All workshop participants disclosed receipt of honoraria and travel reimbursement from the American Thoracic Society.

### References

1. Sackett D. Rules of evidence and clinical recommendations on the use of antithrombotic agents. *Arch Intern Med* 1986;146:464–465.
2. Canadian Task Force on the Periodic Health Examination. The periodic health examination. *CMAJ* 1979;121:1193–1254.
3. Phillips B, Ball C, Sackett D, Badenoch D, Straus S, Haynes B, Dawes M. Levels of evidence and grades of recommendation. Centre for Evidence-Based Medicine (accessed October 29, 2012). Available from: http://www.cebm.net/?o=1025.
4. Fishman A, Martinez F, Naunheim K, Piantadosi S, Wise R, Ries A, Weinmann G, Wood DE. A randomized trial comparing lung-volume-reduction surgery with medical therapy for severe emphysema. *N Engl J Med* 2003;348:2059–2073.
5. Global Strategy for the Diagnosis, Management and Prevention of COPD (GOLD) 2010. 2010 [accessed 2011 Sep 2]. Available from: http://www.goldcopd.org/uploads/users/files/GOLDReport_April112011.pdf.
6. Schunemann HJ, Woodhead M, Anzueto A, Buist S, Macnee W, Rabe KF, *et al*. A vision statement on guideline development for respiratory disease: the example of COPD. *Lancet* 2009;373:774–779.
7. Guyatt G, Akl E, Oxman A, Wilson K, Puhan MA, Wilt T, Gutterman D, Woodhead M, Antman E, Schünemann HJ. Synthesis, grading and presentation of evidence in guidelines. Article 7 in integrating and coordinating efforts in COPD guideline development. An official ATS/ERS workshop report. *Proc Am Thorac Soc* 2012;9:256–261.
8. Wilson K, Irwin RS, File TM, Schunemann HJ, Guyatt G, Rabe KF. Reporting and publishing guidelines. Article 12 in integrating and coordinating efforts in COPD guideline development. An official ATS/ERS workshop report. *Proc Am Thorac Soc* 2012;9:293–297.
9. Wilt T, Guyatt G, Kunz R, MacNee W, Puhan MA, Viegi G, Woodhead M, Schünemann HJ. Deciding what kind of evidence to include in guidelines. Article 5 in integrating and coordinating efforts in COPD guideline development. An official ATS/ERS workshop report. *Proc Am Thorac Soc* 2012;9:243–250.
10. Schunemann HJ, Fretheim A, Oxman AD. Improving the use of research evidence in guideline development: 10. Integrating values and consumer involvement. *Health Res Policy Syst* 2006;4:22.
11. Schunemann HJ, Fretheim A, Oxman AD. Improving the use of research evidence in guideline development: 9. Grading evidence and recommendations. *Health Res Policy Syst* 2006;4:21.
12. DesRoches CM, Campbell EG, Vogeli C, Zheng J, Rao SR, Shields AE, *et al*. Electronic health records' limited successes suggest more targeted uses. *Health Aff (Millwood)* 2010;29:639–646.

13. Shukla V, Bai A, Milne S, Wells G. Systematic review of the evidence grading system for grading level of evidence. *Ger J Evid Qual Health Care* 2008;102(Suppl VI):43.
14. Atkins D, Best D, Briss PA, Eccles M, Falck-Ytter Y, Flottorp S, *et al*. Grading quality of evidence and strength of recommendations. *BMJ* 2004;328:1490.
15. Atkins D, Briss PA, Eccles M, Flottorp S, Guyatt GH, Harbour RT, *et al*. Systems for grading the quality of evidence and the strength of recommendations II: pilot study of a new system. *BMC Health Serv Res* 2005;5:25.
16. Atkins D, Eccles M, Flottorp S, Guyatt GH, Henry D, Hill S, *et al*. Systems for grading the quality of evidence and the strength of recommendations I: critical appraisal of existing approaches. The GRADE Working Group. *BMC Health Serv Res* 2004;4:38.
17. Guyatt G, Gutterman D, Baumann M, Addrizzo-Harris D, Hylek E, Phillips B, Raskob G, Lewis S, Schünemann H. Grading strength of recommendations and quality of evidence in clinical guidelines. *Chest* 2006;129:174–181.
18. Guyatt G, Vist G, Falck-Ytter Y, Kunz R, Magrini N, Schunemann H. An emerging consensus on grading recommendations? [editorial] *ACP J Club* 2006;144:A08.
19. Schünemann HJ, Best D, Vist G, Oxman AD. Letters, numbers, symbols and words: how to communicate grades of evidence and recommendations. *CMAJ* 2003;169:677–680.
20. Institute of Medicine. Current best practices and standards for development of trustworthy CPGs: Part II, Traversing the process. In: Graham R, Mancher M, Miller Wolman D, Greenfield S, Steinberg E, editors. Clinical practice guidelines we can trust. Washington, DC: The National Academies Press; 2011.
21. Fleisher LA, Bass EB, McKeown P. Methodological approach: American College of Chest Physicians guidelines for the prevention and management of postoperative atrial fibrillation after cardiac surgery. *Chest* 2005;128:17S–23S.
22. Oxman AD, Schunemann HJ, Fretheim A. Improving the use of research evidence in guideline development: 14. Reporting guidelines. *Health Res Policy Syst* 2006;4:26.
23. Oxman AD, Schunemann HJ, Fretheim A. Improving the use of research evidence in guideline development: 8. Synthesis and presentation of evidence. *Health Res Policy Syst* 2006;4:20.
24. Steurer-Stey C, Bachmann LM, Steurer J, Tramèr MR. Oral purified bacterial extracts in chronic bronchitis and COPD: systematic review. *Chest* 2004;126:1645–1655.
25. Sin DD, McAlister FA, Man SF, Anthonisen NR. Contemporary management of chronic obstructive pulmonary disease: scientific review. *JAMA* 2003;290:2301–2312.
26. Puhan MA, Garcia-Aymerich J, Frey M, ter Riet G, Anto JM, Agusti AG, *et al*. Expansion of the prognostic assessment of patients with chronic obstructive pulmonary disease: the updated BODE index and the ADO index. *Lancet* 2009;374:704–711.
27. Schunemann H. From BODE to ADO to outcomes in multimorbid COPD patients. *Lancet* 2009;374:667–668.
28. Ai-Ping C, Lee KH, Lim TK. In-hospital and 5-year mortality of patients treated in the ICU for acute exacerbation of COPD: a retrospective study. *Chest* 2005;128:518–524.
29. Qaseem A, Snow V, Owens DK, Shekelle P. The development of clinical practice guidelines and guidance statements of the American College of Physicians: summary of methods. *Ann Intern Med* 2010;153:194–199.
30. Hill J, Bullock I, Alderson P. A summary of the methods that the National Clinical Guideline Centre uses to produce clinical guidelines for the National Institute for Health and Clinical Excellence. *Ann Intern Med* 2011;154:752–757.
31. Guyatt GH, Oxman AD, Kunz R, Falck-Ytter Y, Vist GE, Liberati A, *et al*. Going from evidence to recommendations. *BMJ* 2008;336:1049–1051.
32. Schunemann HJ, Hill SR, Kakad M, Vist GE, Bellamy R, Stockman L, *et al*. Transparent development of the WHO rapid advice guidelines. *PLoS Med* 2007;4:e119.
33. Sawaya GF, Guirguis-Blake J, LeFevre M, Harris R, Petitti D. Update on the methods of the US Preventive Services Task Force: estimating certainty and magnitude of net benefit. *Ann Intern Med* 2007;147:871–875.
34. Hillier S, Grimmer-Somers K, Merlin T, Middleton P, Salisbury J, Tooher R, Weston A. FORM: an Australian method for formulating and grading recommendations in evidence-based clinical guidelines. *BMC Med Res Methodol* 2011;11:23.
35. Morrison JC, Magann EF, Paulson KA, Chauhan SP, Berghella V, Siddiqui D. Evidence-based medicine guidelines in obstetrics/gynecology and trauma surgery. *J Miss State Med Assoc* 2009;50:302–305.
36. Steelman VM, Pape T, King CA, Graling P, Gaberson KB. Selection of a method to rate the strength of scientific evidence for AORN recommendations. *AORN J* 2011;93:433–444.
37. Akl EA, Maroun N, Guyatt G, Oxman AD, Alonso-Coello P, Vist GE, *et al*. Symbols were superior to numbers for presenting strength of recommendations to health care consumers: a randomized trial. *J Clin Epidemiol* 2007;60:1298–1305.
38. Rosenbaum SE, Glenton C, Nylund HK, Oxman AD. User testing and stakeholder feedback contributed to the development of understandable and useful Summary of Findings tables for Cochrane reviews. *J Clin Epidemiol* 2010;63:607–619.
39. Rosenbaum SE, Glenton C, Oxman AD. Summary-of-findings tables in Cochrane reviews improved understanding and rapid retrieval of key information. *J Clin Epidemiol* 2010;63:620–626.
40. Ginnelly L, Claxton K, Sculpher MJ, Golder S. Using value of information analysis to inform publicly funded research priorities. *Appl Health Econ Health Policy* 2005;4:37–46.
41. Claxton K, Cohen JT, Neumann PJ. When is evidence sufficient? *Health Aff (Millwood)* 2005;24:93–101.
42. Brown P, Brunnhuber K, Chalkidou K, Chalmers I, Clarke M, Fenton M, *et al*. How to formulate research recommendations. *BMJ* 2006;333:804–806.
43. Graham ID, Calder LA, Hebert PC, Carter AO, Tetroe JM. A comparison of clinical practice guideline appraisal instruments. *Int J Technol Assess Health Care* 2000;16:1024–1038.
44. Burgers JS, Fervers B, Haugh M, Brouwers M, Browman G, Philip T, *et al*. International assessment of the quality of clinical practice guidelines in oncology using the Appraisal of Guidelines and Research and Evaluation Instrument. *J Clin Oncol* 2004;22:2000–2007.
45. National Guideline Clearinghouse [accessed 2009 Nov 10]. Available from: http://www.guidelines.gov/about/inclusion.aspx
46. Hayward RS, Wilson MC, Tunis SR, Bass EB, Rubin HR, Haynes RB. More informative abstracts of articles describing clinical practice guidelines. *Ann Intern Med* 1993;118:731–737.
47. Burgers J, Anzueto A, Black PN, Cruz AA, Fervers B, Graham ID, Hurd S, Metersky M, Woodhead M, Yawn BP. Adaptation, evaluation and updating of guidelines. Article 14 in integrating and coordinating efforts in COPD guideline development. An official ATS/ERS workshop report. *Proc Am Thorac Soc* 2012;9:304–310.
48. Shekelle PG, Kravitz RL, Beart J, Marger M, Wang M, Lee M. Are nonspecific practice guidelines potentially harmful? A randomized comparison of the effect of nonspecific versus specific guidelines on physician decision making. *Health Serv Res* 2000;34:1429–1448.
49. Hussain T, Michel G, Shiffman RN. The Yale Guideline Recommendation Corpus: a representative sample of the knowledge content of guidelines. *Int J Med Inform* 2009;78:354–363.
50. Lomotan EA, Michel G, Lin Z, Shiffman RN. How "should" we write guideline recommendations? Interpretation of deontic terminology in clinical practice guidelines: survey of the health services community. *Qual Saf Health Care* 2010;19:509–513.
51. Akl E, Guyatt GH, Levine M, Feldstein D, Irani J, Shaneyfelt T, Shaw E, Wasi P, Schünemann HJ. "Might" or "suggest"? No wording approach was clearly superior in conveying the strength of recommendation. *J Clin Epidemiol* 2012;65:268–275.
52. Brozek JL, Bousquet J, Baena-Cagnani CE, Bonini S, Canonica GW, Casale TB, *et al*. Allergic Rhinitis and its Impact on Asthma (ARIA) guidelines: 2010 revision. *J Allergy Clin Immunol* 2010;126:466–476.
53. Chong LY, Nasser M, Glasziou P. GRADE Working Group e-list discussants. What should we call weak recommendations? In: Glasziou P, Guyatt G, editors. *Newsletter of the International Society for Evidence-Based Health Care*, Newsletter 2, January 2011, 6–7.
54. Ahmed F, Temte JL, Campos-Outcalt D, Schunemann HJ. Methods for developing evidence-based recommendations by the Advisory Committee on Immunization Practices (ACIP) of the US Centers for Disease Control and Prevention (CDC). *Vaccine* 2011;29:9171–9176.

# Reporting and Publishing Guidelines

## Article 12 in Integrating and Coordinating Efforts in COPD Guideline Development. An Official ATS/ERS Workshop Report

**Kevin C. Wilson, Richard S. Irwin, Thomas M. File, Jr., Holger J. Schünemann, Gordon H. Guyatt, and Klaus F. Rabe; on behalf of the ATS/ERS Ad Hoc Committee on Integrating and Coordinating Efforts in COPD Guideline Development**

***Introduction*: Professional societies, like many other organizations around the world, have recognized the need to use rigorous processes to ensure that health care recommendations are informed by the best available research evidence. This is the twelfth of a series of 14 articles that were prepared to advise guideline developers in respiratory and other diseases. This article discusses the reporting and publishing of guidelines.**

***Methods*: The authors formulated and discussed the following questions on the reporting and publishing of guidelines. (*1*) What should be reported in guidelines? (*2*) How should guidelines be written? (*3*) How should the bottom-line message be conveyed? (*4*) How should guidelines be packaged? (*5*) Where should guidelines be published? (*6*) Who benefits from the publication of guidelines? (*7*) What information should be vetted by the editor(s)? (*8*) How should guidelines be peer reviewed? We conducted a review of the literature, looking for systematic reviews and methodological research that addressed these questions, but we did not conduct a full systematic review. Our conclusions are based on the available evidence from the published literature and logical arguments from experienced guideline developers.**

***Results and Discussion*: There is little empirical evidence that addresses the reporting and publishing of guidelines. A standard format for reporting guidelines is desirable to ensure that guidelines are comprehensive and that all of the information necessary to judge their quality is presented. In addition, guidelines should contain concise evidence-based recommendations. To facilitate the use of guidelines by consumers, it is preferable to publish them in journals that serve the target audience and to package them in multiple ways. Editors and peer reviewers should ensure that reporting standards have been met, potential conflicts of interest have been adequately addressed and made public, and that the recommendations address important clinical questions.**

## INTRODUCTION

Professional societies, like many other organizations around the world, have recognized the need to use more rigorous processes to ensure that health care recommendations are informed by the best available research evidence. In June 2007 the American Thoracic Society (ATS) and the European Respiratory Society (ERS) convened an international workshop of methodologists and researchers from around the world to coordinate efforts in guideline development, using chronic obstructive pulmonary disease (COPD) as a model (1).

This is the twelfth of a series of 14 articles that were prepared to advise those who develop guidelines related to respiratory and other diseases. The focus of this article is the reporting and publishing of guidelines, which are essential steps in the dissemination of recommendations to the health care community.

## METHODS

The authors formulated and discussed the questions shown in Table 1. We did not conduct a full systematic review of the literature according to standard methodology. To find evidence relevant to the questions, we searched PubMed, the Cochrane Methodology Register (2), and the National Guideline Clearinghouse (3) for systematic reviews and methodological research.

For PubMed, we used the MeSH database with the following search terms: "Publishing [MH] AND Guideline [TIAB]," "Peer Review, Research [MH] AND Guideline [TIAB]," "Reporting [TW] AND Guideline [TIAB]," "Format [TW] AND Guideline [TIAB]," "Standards [TW] AND Guideline [TIAB]," and "Documentation [TW] AND Guideline [TIAB]." The Cochrane Methodology Register was searched by combining the abstract words format, reporting, or publishing with the key words, "CMR: applicability and recommendations—recommendations" or "CMR: applicability and recommendations—levels of evidence and strength of recommendations." The National Guideline Clearinghouse was searched using the annotated bibliography. Search terms included reported, publishing, format, structure, standards, and documentations. Finally, related articles and the reference list from retrieved articles were also reviewed. The literature search has been updated periodically since the workshop and is current through May of 2011.

Our conclusions are based on evidence from the published literature, workshop discussion, logical arguments, and consensus from experienced guideline developers and editors who participated in the writing of this article.

## RESULTS

Database searches did not yield any systematic reviews related to the reporting or publishing of guidelines. Several observational studies, conference proceedings, and a prior review on this topic were identified (4). The important findings are described below.

### 1. What Should Be Reported in Guidelines?

During the Conference on Guideline Standardization (COGS), representatives from 22 organizations with expertise in developing and publishing clinical practice guidelines used a two-stage modified Delphi process to generate reporting standards for clinical practice guidelines (5). The checklist shown in Table 2 lists the 18 topics that participants in the COGS believed should be routinely reported in clinical practice guidelines. Although several

---

This article is a section of "Integrating and Coordinating Efforts in Chronic Obstructive Pulmonary Disease (COPD) Guideline Development," an American Thoracic Society (ATS) and European Respiratory Society (ERS) Workshop Report. This official ATS/ERS Workshop Report was adopted by the ATS Board of Directors, August 2012, and by the ERS Executive Committee, February 2012.

Funding for this conference was made possible (in part) by 1R13HL 90485–01 from the National Institutes of Health. The views expressed in written conference materials or publications and by speakers and moderators do not necessarily reflect the official policies of the Department of Health and Human Services; nor does mention of trade names, commercial practices, or organizations imply endorsement by the U.S. Government.

**Proc Am Thorac Soc Vol 9, Iss. 5, pp 293–297, Dec 15, 2012**

**DOI: 10.1513/pats.201208-065ST**
**Internet address: www.atsjournals.org**

**TABLE 1. QUESTIONS ADDRESSED REGARDING THE REPORTING AND PUBLISHING OF CLINICAL PRACTICE GUIDELINES**

| |
|---|
| Reporting guidelines |
| 1. What should be reported in guidelines? |
| 2. How should guidelines be written? |
| 3. How should the bottom-line message be conveyed? |
| 4. How should guidelines be packaged? |
| Publishing guidelines |
| 5. Where should guidelines be published? |
| 6. Who benefits from the publication of guidelines? |
| 7. What information should be vetted by the editor(s)? |
| 8. How should guidelines be peer reviewed? |

organizations use these reporting standards (5), other organizations and journals have developed their own reporting standards (6) or follow the National Guideline Clearinghouse's Template of Guideline Attributes (3).

We believe it is desirable to have a standard format for reporting guidelines that all organizations that develop guidelines accept. This will ensure that guidelines are comprehensive and that all of the information necessary to judge their quality is presented. Such a format can be generated from existing reporting standards, as well as instruments that are used to evaluate clinical practice guidelines (7–11).

The standard format would ideally require that the following be reported: rationale for the guideline, target audience, scope (patient population, interventions and outcomes considered), methodology (questions asked, literature searches, evidence selection and analysis, formulation and grading of recommendations), recommendations (including clear specification of the population to whom the recommendation applies and the alternatives considered), evidence supporting the recommendations, algorithm(s), references, completion date, guideline committee, authors, conflict of interest disclosures (including how they were reviewed and how conflicts were resolved), funding sources, and sponsoring organizations. With respect to the recommendations, we support the approach advocated by the Institute of Medicine (IOM, Washington, DC), which involves describing and explaining differences of opinion regarding the recommendations, or reporting the results of voting, because it maximizes transparency (11). Although some guideline developers may choose not to include content for every item, they should explicitly address whether the item was considered by the guideline development team (6).

Standardized reporting will also permit assessment of the quality of guidelines with appraisal tools such as the AGREE (Appraisal of Guidelines, Research and Evaluation) II tool (9).

### 2. How Should Guidelines Be Written?

For guidelines to be helpful to clinicians at the point of care, they need to be concise so that information can be located easily and quickly. Excessive time spent trying to find information is a deterrent to seeking the answer to a clinical question (12). The duration judged excessive by most clinicians is uncertain, but is probably quite short. In one observational study, the median time spent pursuing the answer to a clinical question was less than 2 minutes (13).

Guidelines also need to be comprehensive because most are used by a variety of stakeholders (14). Organization of guidelines into many sections, each with a descriptive header and predictable writing style (i.e., the most important information is consistently stated at the beginning or end of a section), is an approach that allows guidelines to be both comprehensive and concise. The descriptive headers make it easier for readers to locate the section that contains the answer to their clinical question, and a predictable writing style makes it easier to locate pertinent information within each section.

Presenting information in the same sequence as a typical patient encounter (e.g., diagnosis before treatment) improves clinician understanding (15). Lists, tables, bolded subheadings, and algorithms are preferred over lengthy uninterrupted prose (16).

### 3. How Should the Bottom-Line Message Be Conveyed?

For each subject addressed in a clinical practice guideline, it is desirable to convey the bottom-line message in the form of one or more recommendations. Although there are no consensus standards for how recommendations should be formulated (6), an optimal recommendation indicates the population for which

**TABLE 2. EIGHTEEN TOPICS THAT PARTICIPANTS IN THE CONFERENCE ON GUIDELINE STANDARDIZATION FELT SHOULD BE ROUTINELY REPORTED IN CLINICAL PRACTICE GUIDELINES***

| Topic | Description |
|---|---|
| 1. Overview | Structure the abstract to include the guideline's release date, status (original, revised, updated), and print and electronic sources |
| 2. Focus | Describe the primary disease and intervention that the guideline addresses, as well as alternative interventions that were considered |
| 3. Goal | Describe the rationale for developing the guideline and the goal that following the guideline is expected to achieve |
| 4. Users | Describe the intended users of the guideline (e.g., type of clinician, patients) and the settings in which use of the guideline is intended |
| 5. Population | Describe the patient population to which recommendations may be applied, as well as exclusion criteria |
| 6. Developer | Identify the organizations responsible for the guideline's development, as well as the names, credentials, and potential conflicts of interest of the participants |
| 7. Funding | Identify the funding source and its role in developing or reporting the guideline, including potential conflicts of interest |
| 8. Evidence | Describe the methods used to search the scientific literature, including the range of dates, databases searched, and criteria used to filter the retrieved evidence |
| 9. Grading | Describe the criteria used to rate the quality of the evidence that supports the recommendations and the system for describing the strength of the recommendation |
| 10. Synthesis | Describe how the evidence was used to develop recommendations (e.g., meta-analysis, evidence tables) |
| 11. Review | Describe how the guideline developer reviewed and tested the guideline before its release |
| 12. Update | Indicate whether the guideline will be updated and if there is an expiration date for the current version |
| 13. Definitions | Define terms that may be unfamiliar or subject to misinterpretation |
| 14. Recommendations | State recommendations precisely and support each with evidence. Indicate the strength of the recommendation and the quality of the evidence |
| 15. Benefits/harms | Describe the anticipated benefits and potential harms associated with implementation of the guideline recommendations |
| 16. Preferences | Describe the role of patient preferences when the recommendation involves substantial personal choice |
| 17. Algorithm | Illustrate the steps in clinical care described by the guideline, if appropriate |
| 18. Implementation | Describe potential barriers to the application of the recommendations |

From Reference 5.

the recommendation is intended, the intervention that is being recommended, and the alternative approach or intervention (17). As an example, a reasonable recommendation may state, "We recommend that all patients with COPD use a short-acting β-agonist on an as-needed basis, rather than on a regularly scheduled basis."

Most organizations that develop guidelines grade the quality of the evidence and the strength of the recommendations. Many groups have adopted the Grading of Recommendations, Assessment, Development, and Evaluation (GRADE) system of grading recommendations (6, 18–20). The formulation and grading of recommendations using the GRADE system are discussed in separate articles within this series (17, 18).

### 4. How Should Guidelines Be Packaged?

Most clinical practice guidelines are reported as full guidelines, which include references and notes (6, 21). Many are also distributed as an executive summary, a summary of recommendations, separate versions for different users, and/or application tools (e.g., algorithms) (6).

We support the packaging of guidelines in multiple ways because this facilitates their use by different consumers. Busy clinicians likely rely on summaries and application tools at the point of care, whereas organizations are more likely to benefit from the more comprehensive full guidelines. When summary versions are reported, the recommendations should contain sufficient information that they can be understood without reference to other supporting material (22).

Evidence summaries, such as full evidence profiles or summary of findings tables (4, 23), should accompany the guideline. Such summaries improve the understanding of the evidence and shorten the duration needed to retrieve pertinent information from systematic reviews (24). The evidence summaries should be transparently linked to the related recommendation, thereby facilitating understanding and judgments about whether the evidence applies to the health care situation. Extensive web repositories of the evidence that influences recommendations are a feasible and attractive option for complete reporting.

Clinical practice guidelines developed by national organizations should be written and packaged in a way that allows them to be modified for local use. As an example, recommending a class of medications (e.g., long-acting β-agonist) in a guideline allows a local institution to apply the recommendation, using a medication that exists in its formulary (e.g., salmeterol or formoterol).

Electronic publication facilitates the packaging of guidelines in multiple formats. An electronically published guideline can provide hyperlinks that quickly direct readers to an executive summary, a summary of recommendations, application tools, alternative versions, or other derivates. Each of these formats can similarly provide hyperlinks that direct readers to the primary version of the guideline or other derivatives, thereby creating a web of hyperlinked versions of the guideline. For printed guidelines, it may be helpful to provide a section that lists the universal resource locators (URLs, or web addresses) of the electronic version of the guideline and its derivatives.

### 5. Where Should Guidelines Be Published?

Clinical practice guidelines that are effectively disseminated and implemented can significantly impact clinical practice (25, 26). In contrast, guidelines will have only a small influence on clinical practice if they are published without a strategy to promote implementation, even if they are published in a prominent medical journal, address an important topic, are of high quality, and are frequently cited (25, 27). A separate article within this series addresses guideline dissemination and implementation (28).

Despite the lack of evidence that publication site matters, we believe that it is desirable to publish a guideline in a medical journal that serves a guideline's target audience. The journal that publishes the guideline is often supported by the same organization that funded the development of the guideline. When guideline development is funded by two or more organizations, the target journal should be determined in advance and documented in a memorandum of understanding between the organizations.

A guideline may occasionally target a heterogeneous audience that is served by different journals. As an example, a guideline about pneumonia may target the pulmonary and infectious disease communities, which tend to read different journals. In such cases, maximizing visibility of the guideline may warrant duplicate publication in different journals, but this needs to be agreed on by the editors of the journals and/or the sponsoring organizations. Duplicate publication is the exception rather than the norm because it can be expensive, difficult to coordinate publication and updating, and may involve copyright issues (29). However, it facilitates dissemination. When an agreement is reached to allow duplicate publication, some organizations choose to publish only the executive summary to decrease cost. Publication in journals that are translated into multiple languages is another way to improve the visibility of a guideline (30).

Medical journals frequently have different policies regarding electronic open access. In particular, the duration before an article can be accessed electronically without cost to the reader varies among journals. Many funding agencies require that any work that they support be submitted to journals with open access policies that are acceptable to the funding agency. As an example, the Medical Research Council of the United Kingdom requires that any work they support be published in journals that provide open access within 6 months of original publication (31). Such policies may affect dissemination of a guideline and should be considered carefully when the funding source and target journal are chosen.

### 6. Who Benefits from the Publication of Guidelines?

Publication of clinical practice guidelines benefits many individuals and organizations. Journals that publish guidelines benefit because, like review articles, guidelines tend to be frequently cited and will improve a journal's impact factor (32). The impact factor is a measure of the frequency with which a particular medical or scientific journal is referenced; it has become a measure of the importance of a journal. Although data are lacking, it is reasonable to expect that a higher impact factor may improve a journal's prestige and the quality of its submissions. This may increase readership, subscriptions, and the profits of the professional society or organization that owns the journal.

Journal editors may also benefit from the publication of guidelines because they may receive credit for improving the journal's standing within the medical community. In addition, authors generally receive academic credit and patients may receive better medical care (25).

### 7. What Information Should Be Vetted by the Editor(s)?

There is little evidence regarding the role of journal editors in the publication of clinical practice guidelines. However, the panel drew several conclusions based on logical arguments.

Guideline development frequently depends on industry funding, committee members often have financial relationships with industry, and journals usually sell reprints of guidelines to commercial

entities. These arrangements are potential financial conflicts of interest. Conflicts of interest may also be intellectual. An intellectual conflict of interest exists when committee members are involved in activities that create a vested interest in a particular viewpoint that is so strongly held that it affects their ability to consider alternative beliefs. Intellectual conflicts of interest are more difficult to identify than financial conflicts. Publicly stated viewpoints, authorship in key trials, and peer-reviewed grant funding are features of a committee member's profile that may prompt additional scrutiny for potential intellectual conflicts of interest (33). Both financial and intellectual conflicts have the potential to introduce bias into the development of recommendations that readers expect to be impartial and evidence-based.

Potential conflicts of interest should be identified and addressed before guideline development commences. An oversight committee may be helpful in facilitating this process (34). Strategies that have been used to minimize the impact of competing interests include insisting that industry funding be in the form of an unrestricted grant with the commercial entity having no input into the process, limiting participants with significant conflicts of interest, excluding participants with significant conflicts of interest, and/or having a methodologist without important conflicts bear the primary responsibility for the final presentation of the evidence. It is important to have transparent disclosure, but disclosure alone is not enough to avoid bias.

The editor(s) of a journal have the unique opportunity and duty to review both the final guideline and the disclosures of the participants. They and the peer reviewers are often the final judges of whether a guideline contains real or perceived conflicts of interest. The editor(s) may be the final arbitrator when disagreement arises and should have the authority to require modification as a condition for publication. A separate article in this series addresses guideline funding and conflicts of interest (35).

The editor(s) must also determine whether the guideline complies with the journal's reporting standards. Such standards ensure comprehensive reporting and that all of the information necessary to judge the quality of a guideline is included. Reporting standards for guidelines are discussed in a previous section (What Should Be Reported in Guidelines).

### 8. How Should Guidelines Be Peer Reviewed?

The IOM has released standards for developing clinical practice guidelines, which include peer review (11). The standards state that peer reviewers should comprise a full spectrum of relevant stakeholders (e.g., scientific and clinical experts, organizations, patients, and the public) and that a draft of the guideline should be made available to the general public for comment at the peer review stage or immediately after it. In addition, peer reviewers' comments should be kept confidential and a record should be kept of the rationale for modifying or not modifying the guideline in response to each comment.

Adherence with the IOM standards occurs to various degrees among organizations that develop guidelines. The National Heart, Lung, and Blood Institute (NHLBI) of the National Institutes of Health (Bethesda, MD) employs a two-step peer review process (36). Clinical practice guidelines are initially reviewed by individuals with expertise similar to that of the guideline panel, typically researchers and methodologists. Intended users are also included in this first round of peer review. The purpose of this step of peer review is to review the draft for accuracy, practicality, clarity, organization, and usefulness of the recommendations. The draft is then modified and subjected to the second step of peer review. The revised draft is posted on the NHLBI website and a call is issued for public review and comment. Comments can be posted on the NHLBI website or made during a public forum that is conducted by the guideline committee.

The National Institute for Health and Clinical Excellence (NICE, London, UK) also uses a multistep peer review process (37, 38). Guidelines are constructed by the Guideline Development Group of NICE, which includes a variety of stakeholders (e.g., physicians, nurses, ancillary personnel, patients). The Consensus Reference Group of NICE then votes on the resulting recommendations. Once the recommendations are agreed on, the guideline is posted on the NICE website and sent to a variety of stakeholders for peer review, including professional organizations. Comments are then discussed by the Guideline Development Group and the guideline is modified. The revised guideline may be finalized or may be subjected to another round of peer review. An independent Guideline Review Panel validates the final full guideline after checking to make sure that stakeholder comments have been taken into account.

The ATS, ERS, American College of Chest Physicians, Infectious Diseases Society of America, and possibly other organizations, have guideline editors and committees that facilitate peer review and help with adherence to methodological criteria in guideline development.

The strength of these approaches is that all stakeholders have the opportunity to comment on the quality of a guideline. Potential pitfalls are that some journals do not routinely perform peer review when a guideline is submitted because the peer review has already occurred as part of a separate document development process, and the multiple steps may delay the release, dissemination, and implementation of the guideline while seemingly minor disagreements are resolved. In addition, the accessibility of the document during the later stages of the peer review process makes it possible for the guideline to be applied to clinical practice before it is finalized and may conflict with a journal's embargo policy. This conflict may influence the journal that is selected to publish a guideline, cause a journal to make an exception to its embargo policy and lose revenue, or lead a guideline to not adhere to the IOM's standards and forego public comment to adhere to the embargo policy.

## CONCLUSIONS

There is limited empirical evidence about the reporting and publishing of clinical practice guidelines. As a result, many of our conclusions are based on logical arguments from experienced guideline developers.

A standard format for reporting guidelines is desirable to ensure that guidelines are comprehensive and that all of the information necessary to judge their quality is presented. In addition, guidelines should contain succinct evidence-based recommendations. We support packaging guidelines in multiple ways to facilitate their use by various consumers, whenever resources permit. This includes full guidelines, an executive summary, summary of recommendations, separate versions for different users, and/or application tools, such as algorithms.

It is preferable to publish a guideline in a journal that maximizes its visibility to the target audience. Editor(s) and peer reviewers should rigorously vet guidelines to ensure that reporting standards have been met and potential conflicts of interest have been adequately addressed. It is appropriate to involve methodologists, researchers, clinicians, and other stakeholders in the peer review of guidelines.

***Author Disclosures***: K.C.W. was employed by UpToDate, Inc. R.S.I. reported no commercial interests relevant to subject matter. T.M.F. was on advisory committees and received lecture fees from Bayer, GlaxoSmithKline, Merck, Ortho-McNeil, Pfizer, Sanofi-Aventis, Schering and Wyeth. He received grants from Bayer, Binax, Ortho-McNeil, Oscient, Pfizer, Sanofi-Aventis and Wyeth. H.J.S. was a member of the GRADE Working Group, which received honoraria and lecture fees regarding

GRADE. G.H.G. was a member of the GRADE Working Group. K.F.R. reported no commercial interests relevant to subject matter. All workshop participants disclosed receipt of honoraria and travel reimbursement from the American Thoracic Society.

## References


1. Schünemann HJ, Woodhead M, Anzueto A, Buist S, Macnee W, Rabe KF, Heffner J. A vision statement on guideline development for respiratory disease: the example of COPD. *Lancet* 2009;373:774–779.
2. Cochrane Methodology Register [Internet] [accessed September 2012]. Available from: http://www.cochrane.org/
3. National Guideline Clearinghouse [Internet] [accessed September 2012]. Available from: http://www.guidelines.gov
4. Schünemann H, Oxman AD, Higgins JPT, Vist GE, Glasziou P, Guyatt GH. Chapter 11: Presenting results and "Summary of findings" tables. In: Higgins JPT, Green S, editors. Cochrane handbook for systematic reviews of interventions, version 501; The Cochrane Collaboration, 2008 [Internet] [accessed September 2012]. Available from: www.cochrane-handbook.org
5. Shiffman RN, Shekelle P, Overhage JM, Slutsky J, Grimshaw J, Deshpande AM. Standardized reporting of clinical practice guidelines: a proposal from the conference on guideline standardization. *Ann Intern Med* 2003;139:493.
6. Oxman AD, Schünemann HJ, Fretheim A. Improving the use of research evidence in guideline development. 14. Reporting guidelines. *Health Res Policy Syst* 2006;4:26.
7. Institute of Medicine. Guidelines for clinical practice: from development to use. Washington, DC: National Academy Press; 1992.
8. Shaneyfelt TM, Mayo-Smith MF, Rothwangl J. Are guidelines following guidelines? The methodological quality of clinical practice guidelines in the peer reviewed medical literature. *JAMA* 1999;281:1900.
9. Brouwers MC, Kho ME, Browman GP, Burgers JS, Cluzeau F, Feder G, Fervers F, Graham ID, Grimshaw J, Hanna SE, *et al*. AGREE II: advancing guideline development, reporting and evaluation in health care. *CMAJ* 2010;182:E839–E842.
10. Graham ID, Calder LA, Hebert PC. A comparison of clinical practice guideline appraisal instruments. *Int J Technol Assess Health Care* 2000;16:1024.
11. Graham R, Mancher M, Wolman DM, Greenfield S, Steinberg E. Clinical practice guidelines we can trust. Washington, DC: National Academies Press; 2011.
12. Ely JW, Osheroff JA, Ebell MH, Chambliss ML, Vinson DC, Stevermer JJ, Pifer EA. Obstacles to answering doctors' questions about patient care with evidence: qualitative study. *BMJ* 2002;324:710.
13. Ely JW, Osheroff JA, Ebell MH, Bergus GR, Levy BT, Chambliss ML, Evans ER. Analysis of questions asked by family doctors regarding patient care. *BMJ* 1999;319:358.
14. Cluzeau F, Wedzicha JA, Kelson M, Corn J, Kunz R, Walsh J, Schünemann HJ. Stakeholder involvement: how to do it right. Article 9 in Integrating and coordinating efforts in COPD guideline development. An Official ATS/ERS Workshop Report. *Proc Am Thorac Soc* 2012;9:269–273.
15. Bergus GR, Chapman GB, Levy BT. Clinical diagnosis and the order of information. *Med Decis Making* 1998;18:412.
16. Ely JW, Osheroff JA, Chambliss ML, Ebell MH, Rosenbaum ME. Answering physicians' clinical questions: obstacles and potential solutions. *J Am Med Inform Assoc* 2005;12:217.
17. Schünemann HJ, Oxman A, Akl EA, Brozek J, Montori V, Heffner J, Hill S, Woodhead M, Campos-Outcalt D, Alderson P, *et al*. Moving from evidence to developing recommendations in guidelines. Article 11 in Integrating and coordinating efforts in COPD guideline development. An Official ATS/ERS Workshop Report. *Proc Am Thorac Soc* 2012;9:XXX.
18. Guyatt GH, Akl E, Oxman A, Wilson K, Puhan MA, Wilt T, Gutterman D, Woodhead M, Antman EM, Schünemann HJ. Synthesis, grading, and presentation of evidence in guidelines. Article 7 in Integrating and coordinating efforts in COPD guideline development. An Official ATS/ERS Workshop Report. *Proc Am Thorac Soc* 2012;9:265–261.
19. Guyatt G, Gutterman D, Baumann MH, Addrizzo-Harris D, Hylek EM, Phillips B, Raskob G, Lewis SZ, Schunemann HJ. Grading strength of recommendations and quality of evidence in clinical guidelines. *Chest* 2006;129:174.
20. Guyatt G, Baumann MH, Pauker S, Halperin J, Maurer J, Owens DK, Tosteson AN, Carlin B, Gutterman D, Prins M, *et al*. Addressing resource allocation issues in recommendations from clinical practice guideline panels. *Chest* 2006;129:182.
21. Moynihan R, Oxman AD, Lavis JN, Paulsen E. Evidence-informed health policy: using research to make health systems healthier [a review of organizations that support the use of research evidence in developing guidelines, technology assessments, and health policy for the WHO Advisory Committee on Health Research]. Oslo: Norwegian Knowledge Centre for the Health Services; 2006.
22. National Institutes for Health and Clinical Excellence. Wording the guideline recommendations. National Institutes for Health and Clinical Excellence Handbook. [Internet] [accessed November 2012]. Available from: http://www.nice.org.uk/media/5F2/44/The_guidelines_manual_2009_-_All_chapters.pdf
23. Schünemann HJ, Hill SR, Kakad M, Bellamy R, Uyeki TM, Hayden FG, Yazdanpanah Y, Beigel J, Chotpitayasunondh T, Del Mar C, *et al*. WHO rapid advice guidelines for pharmacological management of sporadic human infection with avian influenza A (H5N1) virus. *Lancet Infect Dis* 2007;7:21–31.
24. Rosenbaum SE, Glenton C, Oxman AD. Evaluation of summary of findings tables for Cochrane reviews: two randomized trials. *J Clin Epidemiol* 2010;63:620–626.
25. Grimshaw JM, Thomas RE, MacLennan G, Fraser C, Ramsay CR, Vale L, Whitty P, Eccles MP, Matowe L, Shirran L, *et al*. Effectiveness and efficiency of guideline dissemination and implementation strategies. *Health Technol Assess* 2004;8:iii-72.
26. Clancy CM, Slutsky JR. Guidelines for guidelines: we've come a long way. *Chest* 2007;132:746.
27. Bero LA, Grilli R, Grimshaw JM, Harvey E, Oxman AD, Thomson MA; Cochrane Effective Practice and Organization of Care Review Group. Closing the gap between research and practice: an overview of systematic reviews of interventions to promote the implementation of research findings. *BMJ* 1998;317:465.
28. Grimshaw JM, Schünemann HJ, Burgers J, Cruz AA, Heffner J, Metersky M, Cook D. Disseminating and implementing guidelines. Article 13 in Integrating and coordinating efforts in COPD guideline development. An Official ATS/ERS Workshop Report. *Proc Am Thorac Soc* 2012;9:298–303.
29. Flanagin A. Duplicate publication. In: Iverson C, Christiansen S, editors. AMA manual of style: a guide for authors and editors, 10th ed. New York: Oxford University Press; 2007. p. 152.
30. Schünemann HJ, Hill SR, Kakad M, Vist GE, Bellamy R, Stockman L, Wisloff TF, Del Mar C, Hayden F, Uyeki TM, *et al*. Transparent development of the WHO rapid advice guidelines. *PLoS Med* 2007;4:e119.
31. MRC guidance on open access to published research. Medical Research Council [Internet] [accessed September 2012]. Available from: http://www.mrc.ac.uk/PolicyGuidance/EthicsAndGovernance/OpenAccessPublishingandArchiving/MRCGuideforResearchersonOpenAccessPublishing/index.htm
32. Monastersky R. The number that's devouring science. *Chron High Educ* 2005;52:A12.
33. Guyatt G, Akl EA, Hirsh J, Kearon C, Crowther M, Gutterman D, Lewis SZ, Nathanson I, Jaeschke R, Schunemann HJ. The vexing problem of guidelines and conflict of interest: a potential solution. *Ann Intern Med* 2010;152:738–741.
34. Baumann MH, Lewis SZ, Gutterman D. ACCP evidence-based guideline development: a successful and transparent approach addressing conflict of interest, funding, and patient-centered recommendations. *Chest* 2007;132:1015.
35. Boyd EA, Akl EA, Baumann M, Curtis JR, Field M, Jaeschke R, Osborne M, Schünemann HJ. Guideline funding and conflicts of interest. Article 4 in Integrating and coordinating efforts in COPD guideline development. An Official ATS/ERS Workshop Report. *Proc Am Thorac Soc* 2012;9:234–242.
36. National Heart, Lung, and Blood Institute. Who writes and reviews clinical guidelines? About Clinical Practice Guidelines [Internet] [accessed September 2012]. Available from: http://www.nhlbi.nih.gov/guidelines/about.htm#who
37. National Institute for Health and Clinical Excellence. Developing NICE clinical guidelines [Internet] [accessed September 2012]. Available from: http://www.nice.org.uk/aboutnice/howwework/developingniceclinicalguidelines/developing_nice_clinical_guidelines.jsp
38. Wedzicha JA. Peer review and NICE COPD guidelines. *Thorax* 2004;59:183.

# Disseminating and Implementing Guidelines

## Article 13 in Integrating and Coordinating Efforts in COPD Guideline Development. An Official ATS/ERS Workshop Report

**Jeremy M. Grimshaw, Holger J. Schünemann, Jako Burgers, Alvaro A. Cruz, John Heffner, Mark Metersky, and Deborah Cook; on behalf of the ATS/ERS Ad Hoc Committee on Integrating and Coordinating Efforts in COPD Guideline Development**

***Background*: Professional societies, like many other organizations around the world, have recognized the need to use rigorous processes to ensure that health care recommendations are informed by the best available research evidence. This is the thirteenth of a series of 14 articles that were prepared to advise guideline developers in respiratory and other diseases. This article focuses on current concepts and research evidence about how to disseminate and implement guidelines optimally on a national and international level to improve quality of care.**

***Methods*: In this article we address the following questions: What frameworks can aid guideline dissemination and implementation; what are the effects of different guideline dissemination and implementation strategies; and, what is the role of guideline developers in guideline dissemination and implementation? We identified existing systematic reviews and relevant methodological research. Our conclusions are based on evidence from published literature, experience from guideline developers, and workshop discussions.**

***Results and Conclusions*: The Knowledge to Action cycle proposed by Graham and colleagues (*J Contin Educ Health Prof* 2006;26:13–24) provides a useful framework for planning dissemination and implementation activities that emphasize the need for tailored approaches based on an assessment of local barriers. There are a broad range of interventions that are generally effective at improving the uptake of evidence. The best intervention depends on likely barriers, available resources, and other practical considerations. Financial interventions (such as pay for performance) appear to be as effective as other interventions that aim to change professional behavior. Guideline developers who do not have responsibility for guideline implementation in their jurisdiction should support those with responsibility for implementation by considering the "implementability" of their guidelines.**

## INTRODUCTION

Evidence-based guidelines provide recommendations for best practice. However, guidelines are not self-implementing. Developing guidelines and making them available to health care professionals does not ensure their use. There is consistent evidence of gaps between practice guidelines and clinical practice (1). For example, a survey of almost 700 Swiss primary care physicians identified key evidence–practice gaps in the management of chronic obstructive pulmonary disease (COPD). Only 55% used spirometric criteria to define COPD; 62% and 29% immunized patients against influenza and *Pneumococcus*, respectively; and only 28% used pulmonary rehabilitation (2). Thus, a key global challenge shared by guideline developers and health care systems is optimal dissemination and implementation of guidelines, thereby turning their recommendations into action.

In June 2007 the American Thoracic Society (ATS) and the European Respiratory Society (ERS) convened an international workshop of methodologists and researchers from around the world to coordinate efforts in guideline development, using COPD as a model (3). Participants maintained working groups during the subsequent 4 years to develop a series of recommendations. This is the thirteenth of a series of 14 articles prepared to advise guideline developers in respiratory and other diseases, and guideline users across the health care system.

In this article we summarize the current concepts and research evidence about how to disseminate and implement guidelines optimally on a national and international level to improve quality of care. In particular, we focus on the frameworks that can aid guideline dissemination and implementation, and the effects of various guideline dissemination and implementation strategies. We also discuss the role of guideline developers (and other stakeholders) in guideline dissemination and implementation, as well as opportunities for implementation research.

## METHODS

After deciding on the detailed scope and questions of the article (Table 1), we searched MEDLINE and other databases of methodological studies for existing systematic reviews and relevant methodological research through December 2010. We did not conduct systematic reviews ourselves. Our conclusions are based on available evidence from published literature, the experience of guideline developers, and our workshop discussions. The conclusions are also based on a critical review of planned change models undertaken by Graham and colleagues (4, 5) and an updated overview of reviews of professional behavior change interventions undertaken by the Cochrane Effective Practice and Organization of Care group (6).

## RESULTS

### 1. What Frameworks Can Aid Guideline Dissemination and Implementation?

Guideline dissemination and implementation frequently require changing the behavior of health care professionals. Work in this field may benefit from knowledge of social science, including theories of behavior and behavior change (7). Planned change models can be used as frameworks for planning guideline dissemination and implementation. These are sets of logically interrelated

---

This article is a section of "Integrating and Coordinating Efforts in Chronic Obstructive Pulmonary Disease (COPD) Guideline Development," an American Thoracic Society (ATS) and European Respiratory Society (ERS) Workshop Report. This official ATS/ERS Workshop Report was adopted by the ATS Board of Directors, August 2012, and by the ERS Executive Committee, February 2012.

Funding for this conference was made possible (in part) by 1R13HL 90485-01 from the National Institutes of Health. The views expressed in written conference materials or publications and by speakers and moderators do not necessarily reflect the official policies of the Department of Health and Human Services; nor does mention of trade names, commercial practices, or organizations imply endorsement by the U.S. Government.

Jeremy Grimshaw holds a Canada Research Chair in Knowledge Transfer and Uptake. Holger Schünemann holds the Michael Gent Chair in Healthcare Research. John Heffner holds the endowed Garnjobst Chair in Medical Education. Deborah Cook holds a Canada Research Chair in Knowledge Translation in Critical Care Medicine.

**Proc Am Thorac Soc Vol 9, Iss. 5, pp 298–303, Dec 15, 2012**

**DOI: 10.1513/pats.201208-066ST**
**Internet address: www.atsjournals.org**

**TABLE 1. QUESTIONS ADDRESSED REGARDING DISSEMINATION AND IMPLEMENTATION OF GUIDELINES**

1. What frameworks can aid guideline dissemination and implementation?
2. What are the effects of various guideline dissemination and implementation strategies?
3. What is the role of guideline developers in guideline dissemination and implementation?

concepts that explain, in a systematic way, the means by which planned change occurs; predict how various forces in an environment will react in specified change situations; and help planners or change agents control variables that increase or decrease the likelihood of the occurrence of change (8).

One of the challenges faced by nonsocial scientists when attempting to understand theories and models from social sciences is the apparent plethora of models. Graham and colleagues undertook a critical review of planned change models (4, 5). They searched the Internet and multiple bibliographic databases for planned change models published between 1980 and May 2005, identified 31 different planned change models from diverse disciplines, and observed conceptual similarities across the models. On the basis of this, they proposed the Knowledge to Action loop (Figure 1) that highlights central processes relating to knowledge creation, distillation, and use. The central *knowledge creation funnel* represents knowledge generation (primary studies), synthesis (systematic reviews), and development of knowledge tools (clinical practice guidelines, decision aids, policy briefs). They note that "as knowledge moves through the funnel, it becomes more distilled and refined and presumably more useful to stakeholders" (5).

The *action cycle* is based on planned action theories that focus on deliberate engineering of change in health care and other systems and includes the following:

- Identification of a problem that needs addressing
- Identification, review, and selection of the knowledge or research relevant to the problem (e.g., practice guidelines or research findings)
- Adaptation of the identified knowledge or research to the local context
- Assessment of barriers to using the knowledge
- Selection, tailoring, and implementation of interventions to promote the use of knowledge (i.e., implement the change)
- Monitoring knowledge use
- Evaluation of the outcomes of using the knowledge
- Sustaining ongoing knowledge use

Barriers to change can exist at various levels of the health care system and include structural barriers (e.g., lack of resources, financial disincentives), organizational barriers (e.g., inappropriate skill mix, lack of facilities or equipment), peer group barriers (e.g., local standards of care not in line with desired practice), professional–patient interaction barriers (e.g., communication and information-processing issues), and competing priorities (9, 10). There are diverse methods to identify barriers that vary in their formality. Commonly, barriers can be identified by informal consultation with key stakeholders. However, it is desirable to use more formal approaches including observation, one-on-one interviews, focus groups, and interviewer or self-administered surveys. It is unclear which way most accurately identifies barriers to change. Barriers may vary for given resources, across settings, and for different guidelines.

Selection of potential interventions to produce change should be based on the consideration of barriers, mechanisms of action of potential interventions (i.e., what interventions are likely to overcome observed barriers), potential contextual effect modifiers, resources needed to deliver dissemination and implementation strategies, and logistical issues. Few reported studies of dissemination and implementation strategies provide justification for the choice and design of interventions (11). Involvement of the target group in the choice of interventions, based on detailed analysis of barriers to implementing the guideline in practice, could increase the acceptance and effectiveness of strategies (12). Formal intervention mapping (13) (which explicitly links the choice of intervention to identified barriers) forces guideline implementers to be more explicit about the program logic behind the choice and design of their intervention(s).

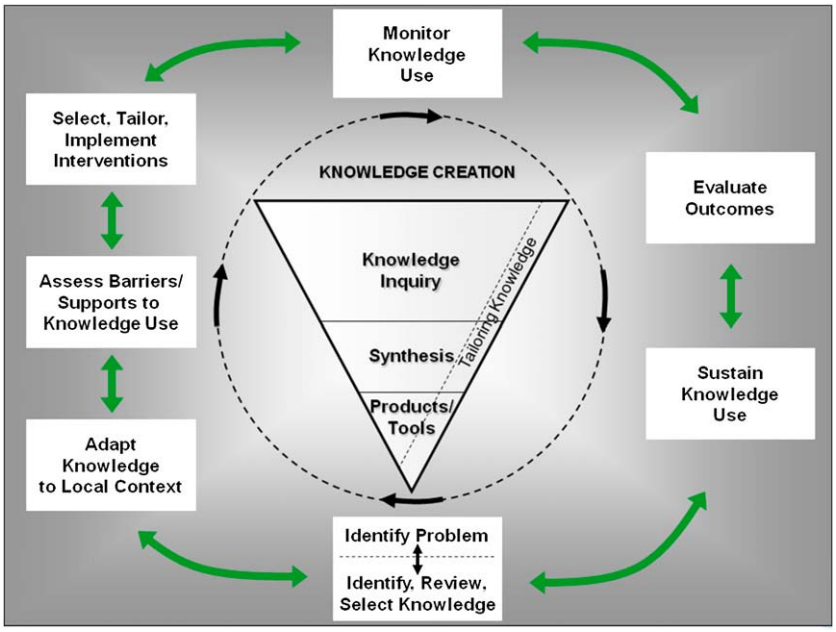


***Figure 1.*** Knowledge to Action cycle (reprinted from Reference 5).

**TABLE 2. SUMMARY OF EFFECTIVENESS OF INTERVENTIONS TO INCREASE GUIDELINE IMPLEMENTATION**

| Intervention (Key Reference) | Definition of Intervention Based on Bero *et al.* (20) | Barriers Addressed | Effectiveness | Resource Considerations | Practical Considerations |
|---|---|---|---|---|---|
| Printed educational materials (21) | Distribution of published or printed recommendations for clinical care (including clinical practice guidelines) | Individual professional knowledge (attitudes) | AMSTAR score, 8<br>6 RCTs<br>Generally effective<br>Median effect size, +4.6% absolute improvement (absolute range, –8.0% to +9.6%) | Relatively inexpensive | Commonly used in health care settings |
| Educational meetings (22) | Health care providers who have participated in conferences, lectures, workshops, or traineeships | Individual professional and peer group knowledge, attitudes, and skills | AMSTAR score, 8<br>81 RCTs<br>Generally effective<br>Median effect size across 36 comparisons, +6% absolute improvement (interquartile range, 1.8% to 15.3%)<br>Larger effects observed with: higher attendance at educational meetings; with mixed interactive and didactic educational meetings; simpler behaviors; and serious outcomes | Relatively inexpensive (didactic) to modest expense (mixed/interactive—usually higher facilitator-to-participant ratio than didactic activities) | Commonly used in health care settings |
| Educational outreach (23) | Use of a trained person who met with providers in their practice settings to give information with the intent of changing the provider's practice | Individual professional knowledge, attitudes using a social marketing approach (24) | AMSTAR score, 8<br>69 RCTs<br>Generally effective<br>Prescribing behaviors: median effect size across 17 comparisons, +4.8% absolute improvement (interquartile range, +3.0% to +6.5%)<br>Other behaviors: median effect across 17 comparisons, +6.0% absolute improvement (interquartile range, +3.6% to +16.0%) | Relatively expensive due to employment of academic detailers (although can still be efficient) (16) | Used in some health care systems. Typically aim to get maximum of 3 messages across in 10–15 min (using approach tailored to individual health care provider) and use additional strategies to reinforce approach (24)<br>Typically focus on relatively simple behaviors in control of individual physician, e.g., choice of drugs to prescribe |
| Local opinion leaders (25) | Use of providers nominated by their colleagues as "educationally influential." The investigators must have explicitly stated that their colleagues identified the opinion leaders | Individual professional and peer group knowledge, attitudes, (skills) | AMSTAR score, 9<br>12 RCTs<br>Generally effective<br>Median effect, 10% absolute improvement (absolute range, –6% to +25%) | Moderately expensive due to need to survey target population for each condition | Rarely used in health care systems. Majority of studies have used the Hiss instrument to identify opinion leaders (who are up-to-date, good communicators, humanistic)<br>Appear to be condition specific (26)<br>Coverage across social networks is often uncertain<br>Sustainability is uncertain (26) |
| Audit and feedback (27) | Any summary of clinical performance of health care over a specified period of time | Individual professional (and peer group) awareness of current performance | AMSTAR score, 8<br>118 RCTs<br>Generally effective<br>Median effect across 88 high-quality comparisons, +5% (interquartile range, +3% to +11%)<br>Greater effects seen if baseline compliance is low | Resources required largely relate to costs of data abstraction. May be relatively inexpensive if data can be abstracted by routine administrative systems | Used in some health care systems. Feasibility may depend on availability of high-quality administrative data |
| Reminders (28) | Patient- or encounter-specific information, provided on a computer screen, which is designed or intended to prompt a health professional to recall information | Individual professional cognitive/memory barriers | AMSTAR score, 8<br>28 RCTs<br>Generally effective<br>Median effect across 32 comparisons, +4.2% (interquartile range, +0.8% to +18.8%) | Resources vary across delivery mechanism. Increasing use of computerized reminders—where inclusion of reminders has relatively modest cost | Used in some health care systems. Insufficient knowledge about how to prioritize and optimize reminders |
| Multifaceted interventions (29) | An intervention including two or more components | Target multiple barriers of the included intervention components | AMSTAR score, 7<br>Grimshaw and colleagues failed to demonstrate a dose–response analysis (i.e., the apparent effects of interventions did not increase with the number of components) | Likely more costly than single interventions | Used in some health care systems. Need to carefully consider how to combine interventions to ensure additive or synergistic effects (e.g., interventions that include components that target same barriers may not be additive/synergistic) |

*Definition of abbreviations*: AMSTAR = Assessment of Multiple Systematic Reviews; RCT = randomized controlled trial.
Derived from Reference 6.

**TABLE 3. "IMPLEMENTABILITY" FRAMEWORK**

| Domain | Element | Examples |
|---|---|---|
| Usability | Navigation | Table of contents |
| | Evidence format | Narrative, tabulated, or both |
| | Recommendation format | Narrative, graphic (algorithms), or both; recommendation summary (single list in full or summary version) |
| Adaptability | Alternative versions | Summary (print, electronic for PDA); patient (tailored for patients/caregivers); published (journal) |
| Validity | Number of references | Total number of distinct references to evidence on which recommendations are based |
| | Evidence graded | A system is used to categorize quality of evidence supporting each recommendation |
| | Number of recommendations | Total number of distinct recommendations (subrecommendations considered same) |
| Applicability | Individualization | Clinical information (indications, criteria, risk factors, drug dosing) that facilitates application of the recommendations explicitly highlighted as tips or practical issues using subtitles or text boxes, or summarized in tables and referred to in recommendations or narrative contextualizing recommendations |
| Communicability | Patient education or involvement | Informational or educational resources for patients/caregivers, questions for clinicians to facilitate discussion, or contact information (phone, fax, e-mail, or URL) to acquire informational or educational resources |
| Accommodation | Objective | Explicitly stated purpose of guideline (clinical decision making, education, policy, quality improvement) |
| | Users | Who would deliver/enable delivery of recommendations (individuals, teams, departments, institutions, managers, policy makers, internal/external agents), who would receive the services (patients/caregivers) |
| | User needs/values | Identification of stakeholder needs, perspectives, interests, or values |
| | Technical | Equipment or technology needed, or the way services should be organized to deliver recommendations |
| | Regulatory | Industrial standards for equipment or technology, or policy regarding their use |
| | Human resources | Type and number of health professionals needed to deliver recommended services |
| | Professional | Education, training or competencies needed by clinicians/staff to deliver recommendations |
| | Impact | Anticipated changes in workflow or processes during/after adoption of recommendations |
| | Costs | Direct or productivity costs incurred as a result of acquiring resources or training needed to accommodate recommendations, or as a result of service reductions during transition from old to new processes |
| Implementation | Barriers/facilitators | Individual, organizational, or system barriers that are associated with adoption |
| | Tools | Instructions, tools or templates to tailor guideline/recommendations for local context; point-of-care templates/forms (clinical assessment, standard orders) |
| | Strategies | Possible mechanisms by which to implement guideline/recommendations |
| Evaluation | Monitoring | Suggestions for evaluating compliance with organization, delivery and outcomes of recommendations, including program evaluation, audit tools, and performance measures/quality indicators |

Reprinted by permission from Reference 19.

### 2. What Are the Effects of Various Guideline Dissemination and Implementation Strategies?

***Professional interventions.*** The Cochrane Effective Practice and Organization of Care (EPOC) group has completed three overviews of professional behavior change interventions (6, 14, 15). On the basis of work to date, summaries of more than 200 systematic reviews of professional behavior change interventions have been made available in the Rx for Change database as of the summer of 2011 (6). Herein, we summarize the effects of key interventions, based on the highest quality reviews in the Rx

for Change database (Table 2) (6). In addition, we consider the likely mechanisms of actions of interventions, the potential barriers that the interventions might address, and practical and logistical issues.

Overall, the results demonstrate that most interventions are effective under certain circumstances, associated with modest but important effects. Although there is a substantial evidence base supporting the effectiveness of some interventions (e.g., audit and feedback, educational outreach), there is much less available evidence about other interventions (e.g., the role of opinion leaders); accordingly, inferences are limited. The resources required to deliver these interventions range from relatively inexpensive (e.g., educational materials) to relatively expensive (e.g., educational outreach). However, given the costs of health care, even relatively small effects of relatively expensive interventions may still be cost-effective (16). Finally, the practical steps in the delivery of interventions are often poorly identified in the available studies.

***Economic interventions.*** There has been considerable interest in the use of financial incentives, such as "pay for performance," to improve quality of care. We summarize the results of the most recently published systematic review of the impact of financial incentives by Petersen and colleagues (17). They identified 17 studies published in English between 1980 and November 2005, including 9 randomized trials showing mixed effects and 4 trials showing statistically significant results. Across the trials, the median absolute improvement was +7% in process of care measures.

Taken together, these reviews highlight that it is possible to change professional behavior and improve the quality of care, using a variety of interventions. Although the effects are often modest, they are potentially important and potentially cost-effective. However, there is considerable variation in the observed effects seen for any intervention, suggesting the importance of barrier identification and intervention specification. The underlying rationale is that greater effects are observed when barriers have been correctly identified and interventions are targeted at these barriers. Addressing barriers commonly also requires system changes to support professional behavior change initiatives.

## 3. What Is the Role of Guideline Developers in Guideline Dissemination and Implementation?

Guideline developers rarely have direct responsibility for guideline dissemination and implementation. For example, in COPD management much of the dissemination is a result of active dissemination by the pharmaceutical industry or professional societies. By contrast, guideline implementation is usually the responsibility of health care practitioners or managers. As a result, it is not always clear what the role of guideline developers should be in dissemination and implementation.

Guideline developers should be concerned with the ability to implement (i.e., the "implementability") their guidelines (18). This implies that developers should develop formal relationships with those in health care systems responsible for guideline dissemination and implementation to support guideline uptake. Gagliardi and colleagues proposed a framework for guideline implementability that identified 22 elements in 8 key domains, including format domains (usability, adaptability, and validity) and content domains (applicability, communicability, accommodation, implementation, and evaluation) (see Table 3) (19). They applied the framework to 20 chronic disease guidelines to determine the extent to which guidelines included elements to improve implementability. They observed that most guidelines "contained a large volume of graded, narrative evidence, and tables featuring complementary clinical information" and that "few contained additional features that could improve guideline use." At present, there is relatively little research evidence about how developers should improve implementability of their guidelines, but there are a number of ongoing studies that should provide guidance for developers (e.g., www.decide-collaboration.eu).

## Guideline Development, Dissemination, and Implementation as Opportunities for Implementation Research

Our current scientific knowledge about how to disseminate and implement guidelines is incomplete. Nevertheless, guideline developers and health care systems are spending considerable resources disseminating and implementing guidelines. As a result, there is a profusion of "natural experiments" of guideline dissemination and implementation that provide considerable opportunities for guideline developers and health care systems to collaborate with implementation researchers to enhance our scientific knowledge about optimal dissemination and implementation approaches. Moreover, findings from implementation research could inform guideline developers in updating guidelines, so that knowledge can effectively be transformed into action.

***Author Disclosures:*** J.M.G. reported no commercial interests relevant to subject matter. H.J.S. was a member of the GRADE Working Group, which received honoraria and lecture fees regarding GRADE. J.B. received lecture fees from Sanofi-Aventis. A.A.C. was employed by the World Health Organization. J.H. reported no commercial interests relevant to subject matter. M.M. consulted with Qualidigm and Pfizer, was on advisory committees of Pfizer, and received lecture fees from Pfizer. D.C. reported no commercial interests relevant to subject matter. All workshop participants disclosed receipt of honoraria and travel reimbursement from the American Thoracic Society.

## References

1. Grol R, Grimshaw J. From best evidence to best practice: effective implementation of change in patients' care. *Lancet* 2003;362:1225–1230.
2. Rutschmann TO, Janssens JP, Vermeulen B, Sarasin FP. Knowledge of guidelines for the management of COPD: a survey of primary care physicians. *Respir Med* 2004;98:932–937.
3. Schunemann HJ, Woodhead M, Anzueto A, Buist S, Macnee W, Rabe KF, Heffner J. A vision statement on guideline development for respiratory disease: the example of COPD. *Lancet* 2009;373:774–779.
4. Graham I, Tetroe J. Some theoretical underpinnings of knowledge translation. *Acad Emerg Med* 2007;14:936–941.
5. Graham ID, Logan J, Harrison MB, Straus SE, Tetroe J, Caswell W, Robinson N. Lost in knowledge translation: time for a map? *J Contin Educ Health Prof* 2006;26:13–24.
6. Canadian Optimal Medication Prescribing and Utilization Service (COMPUS), Canadian Agency for Drugs and Technologies in Health (CADTH). Rx for Change database: interventions directed to professionals, 2012 [Internet] [accessed September 2012]. Available from: http://www.cadth.ca/en/resources/rx-for-change/
7. Eccles M, Grimshaw J, Walker A, Johnston M, Pitts N. Changing the behavior of healthcare professionals: the use of theory in promoting the uptake of research findings. *J Clin Epidemiol* 2005;58:107–112.
8. Rimmer TC, Johnson LL. Planned change theories for nursing: review, analysis and implications. Thousand Oaks, CA: Sage; 1998.
9. Grimshaw J, Eccles M, Tetroe J. Implementing clinical guidelines: current evidence and future implications. *J Contin Educ Health Prof* 2004;24:S31–S37.
10. Cabana MD, Rand CS, Powe NR, Wu AW, Wilson MH, Abboud PA, Rubin HR. Why don't physicians follow clinical practice guidelines? A framework for improvement. *JAMA* 1999;282:1465.
11. Davies P, Walker AE, Grimshaw JM. A systematic review of the use of theory in the design of guideline dissemination and implementation strategies and interpretation of the results of rigorous evaluations. *Implement Sci* 2010;5:14.
12. Lugtenberg M, Zegers-van Schaick JM, Westert GP, Burgers JS. Why don't physicians adhere to guideline recommendations in practice? An analysis of barriers among Dutch general practitioners. *Implement Sci* 2009;4:54.
13. van Bokhoven MA, Kok G, van der Weidjen T. Designing a quality improvement intervention: a systematic approach. *Qual Saf Health Care* 2003;12:215–220.

14. Bero LA, Grilli R, Grimshaw JM, Harvey E, Oxman AD, Thomson MA; Cochrane Effective Practice and Organization of Care Review Group. Closing the gap between research and practice: an overview of systematic reviews of interventions to promote the implementation of research findings. *BMJ* 1998;317:465–468.
15. Grimshaw JM, Shirran L, Thomas R, Mowatt G, Fraser C, Bero L, Grilli R, Harvey EL, Oxman AD, O'Brien MA. Changing provider behavior: an overview of systematic reviews of interventions. *Med Care* 2001;39 (Suppl 2):II2–II45.
16. Mason J, Freemantle N, Nazareth I, Eccles M, Haines A, Drummond M. When is it cost-effective to change the behavior of health professionals? *JAMA* 2001;286:2988–2992.
17. Petersen LA, Woodard LD, Urech T, Daw C, Sookanan S. Does pay-for-performance improve the quality of health care? *Ann Intern Med* 2006;145:265–272.
18. Shiffman RN, Dixon J, Brandt C, Essaihi A, Hsiao A, Michel G, O'Connell R. The GuideLine Implementability Appraisal (GLIA): development of an instrument to identify obstacles to guideline implementation. *BMC Med Inform Decis Mak* 2005;5:23.
19. Gagliardi AR, Brouwers MC, Palda VA, Lemieux-Charles L, Grimshaw JM. How can we improve guideline use? A conceptual framework of implementability. *Implement Sci* 2011;6:26.
20. Grimshaw J, Shepperd S, Ballini L, Bero L, Durieux P, Gruen RL, Leng G, Lewin S, Munabi-Babigumira S, O'Connor D, *et al.*; Cochrane Effective Practice and Organisation of Care Group. About the Cochrane Collaboration. *Cochrane Rev Groups* 2013;1:EPOC.
21. Farmer AP, Legare F, Turcot L, Grimshaw J, Harvey E, McGowan JL, Wolf F. Printed educational materials: effects on professional practice and health care outcomes. *Cochrane Database Syst Rev* [serial on the Internet]. 2008 [accessed September 2012];3:CD004398. Available from: http://summaries.cochrane.org/
22. Forsetlund L, Bjorndal A, Rashidian A, Jamtvedt G, O'Brien MA, Wolf F, Davis D, Odgaard-Jensen J, Oxman AD. Continuing education meetings and workshops: effects on professional practice and health care outcomes. *Cochrane Database Syst Rev* [serial on the Internet]. 2009 [accessed September 2012];2:CD003030. Available from: http://summaries.cochrane.org/
23. O'Brien MA, Rogers S, Jamtvedt G, Oxman A, Odgaard-Jensen J, Kristoffersen D, Forsetlund L, Bainbridge D, Freemantle N, Davis DA, Haynes RB, Harvey EL. Educational outreach visits: effects on professional practice and health care outcomes. *Cochrane Database Syst Rev* [serial on the Internet]. 2007 [accessed September 2012];4:CD000409. Available from: http://summaries.cochrane.org/
24. Soumerai SB, Avorn J. Principles of educational outreach ("academic detailing") to improve clinical decision making. *JAMA* 1990;263: 549–556.
25. Flodgren G, Parmelli E, Doumit G, Gattellari M, O'Brien MA, Grimshaw J, Eccles MP. Local opinion leaders: effects on professional practice and health care outcomes. *Cochrane Database Syst Rev* [serial on the Internet]. 2010 [accessed September 2012];10:CD000125. Available from: http://summaries.cochrane.org/
26. Doumit G. Opinion leaders: Effectiveness, Identification, Stability, Specificity, and Mechanism of Action. Ottawa: University Of Ottawa; 2006.
27. Jamtvedt G, Young JM, Kristoffersen DT, O'Brien MA, Oxman AD. Audit and feedback: effects on professional practice and health care outcomes. *Cochrane Database Syst Rev* [serial on the Internet]. 2006 [accessed September 2012];2:CD000259. Available from: http://summaries.cochrane.org/
28. Shojania KG, Jennings A, Mayhew A, Ramsay CR, Eccles MP, Grimshaw J. The effects of on-screen, point of care computer reminders on processes and outcomes of care. *Cochrane Database Syst Rev* [serial on the Internet]. 2009 [accessed September 2012];3:CD001096. Available from: http://summaries.cochrane.org/
29. Grimshaw JM, Thomas RE, Maclennan G, Fraser C, Ramsay CR, Vale L, Whitty P, Eccles MP, Matowe L, Shirran L, *et al*. Effectiveness and efficiency of guideline dissemination and implementation strategies. *Health Technol Assess* 2004;8:iii-72.

# Adaptation, Evaluation, and Updating of Guidelines

## Article 14 in Integrating and Coordinating Efforts in COPD Guideline Development. An Official ATS/ERS Workshop Report

**Jako S. Burgers, Antonio Anzueto, Peter N. Black, Alvaro A. Cruz, Béatrice Fervers, Ian D. Graham, Mark Metersky, Mark Woodhead, and Barbara P. Yawn; on behalf of the ATS/ERS Ad Hoc Committee on Integrating and Coordinating Efforts in COPD Guideline Development**

***Introduction*: Professional societies, like many other organizations, have recognized the need to use more rigorous processes to ensure that health care recommendations are informed by the best available research evidence. This is the last of a series of 14 articles that methodologists and researchers from around the world have prepared to advise guideline developers in respiratory and other diseases on how to achieve this. We updated a review of the literature on guideline adaptation, evaluation, and updating, focusing on four key questions.**

***Methods*: In this review we addressed the following questions. (*1*) Which high-quality guidelines on chronic obstructive pulmonary disease (COPD) are available? (*2*) How should guidelines be adapted to the user's context and culture? (*3*) How should the use of guidelines be evaluated in clinical practice? and (*4*) How should guidelines be efficiently kept up-to-date? We did not conduct systematic reviews ourselves. We relied on a literature review published in 2006 and on a manual produced by the ADAPTE Collaboration to inform our judgments, as well as our collective experience and workshop discussions.**

***Results and Discussion*: Guideline adaptation can be seen as an alternative to *de novo* development and as part of an implementation process, taking into consideration the user's own context. A systematic approach should be followed to ensure high quality of the resulting guidance. On the topic of COPD, many guidelines are available. Guidelines of the Global Initiative for Chronic Obstructive Lung Disease and of the American Thoracic Society and European Respiratory Society are particularly well-suited for adaptation. The adaptation process includes (*1*) definition of specific questions that need to be answered by the guideline; (*2*) assessment of guideline quality; (*3*) assessment of the clinical content, validity, acceptability, applicability, and transferability of the recommendations; and (*4*) decisions about adoption or adaptation of the recommendations. The use of the guidelines in practice can be measured with performance indicators. Adverse effects of strict adherence to guideline recommendations should be prevented, in particular when the improvement of patient outcomes is unclear. COPD guidelines should be updated at least every 2 years. Collaboration between COPD guideline developers is recommended to prevent duplication of effort.**

## INTRODUCTION

Professional health care societies, like many other organizations around the world, have recognized the need to use more rigorous processes to ensure that health care recommendations are informed by the best available research evidence. This could result in duplication of efforts, if several scientific organizations are active in the same field of reviewing evidence and developing guidelines. Therefore, there is a need for efficient use of resources and international collaboration in guideline development and implementation (1, 2). In this context, the concept of guideline adaptation is receiving increased attention among guideline organizations, health care organizations, and clinical practice settings worldwide (3, 4). Guideline adaptation can be defined as the modification of any existing guidelines for use in a different cultural, regional, or organizational context (5). This includes consideration of language, availability of services, the health care setting, and patients' and providers' cultural and ethical values. Adaptation can be used as an alternative to *de novo* guideline development or for customizing an existing guideline to suit the local context.

## WHAT ARE PROFESSIONAL SOCIETIES DOING NOW?

Until now, there have been no published examples of chronic obstructive pulmonary disease (COPD) guideline adaptation by a professional society. On the other hand, the guideline documents of the Global Initiative for Chronic Obstructive Lung Disease (GOLD) have been translated into several languages (www.goldcopd.org), often demanding cultural as well as language modification of this international guideline for use in different countries and contexts. Similarly, the international consensus statement on COPD sponsored by the American Thoracic Society (ATS) and European Respiratory Society (ERS), available in English for professionals, has been translated into patient information materials in five languages likewise requiring both linguistic and cultural changes (http://www.european-lung-foundation.org/27-european-lung-foundation-elf-copd.htm).

In June 2007 the ATS and the ERS convened an international workshop of methodologists and researchers to consider coordinating efforts in guideline development for COPD and other diseases (6). Participants completed the work during the subsequent 4 years. This is the last of a series of 14 articles resulting from this workshop. In this article we address the issues related to guideline adaptation, evaluation, and updating in the context of respiratory disease guidelines.

## METHODS

We address the following four questions:

1. What high-quality guidelines on COPD are available?
2. How should guidelines be adapted to one's own context and culture?
3. How should the use of guidelines be evaluated in clinical practice?
4. How should guidelines be efficiently kept up-to-date?

To address these questions we used the results of a literature review for the World Health Organization (2) and a review conducted by the ADAPTE Collaboration (5). This group was

---

This article is a section of "Integrating and Coordinating Efforts in Chronic Obstructive Pulmonary Disease (COPD) Guideline Development," an American Thoracic Society (ATS) and European Respiratory Society (ERS) Workshop Report. This official ATS/ERS Workshop Report was adopted by the ATS Board of Directors, August 2012, and by the ERS Executive Committee, February 2012.

Funding for this conference was made possible (in part) by 1R13HL 90485–01 from the National Institutes of Health. The views expressed in written conference materials or publications and by speakers and moderators do not necessarily reflect the official policies of the Department of Health and Human Services; nor does mention of trade names, commercial practices, or organizations imply endorsement by the U.S. Government.

**Proc Am Thorac Soc Vol 9, Iss. 5, pp 304–310, Dec 15, 2012**

**DOI: 10.1513/pats.201208-067ST**
**Internet address: www.atsjournals.org**

initiated during a collaborative project between the French National Federation of Comprehensive Cancer Centers (FNCLCC) and the Quebec Cancer Control Department, with the aim of examining how the FNCLCC guidelines could be adapted for use in French-speaking Quebec. Databases searched included PubMed (from 1966 through March 2011), the U.S. National Guideline Clearinghouse (www.guideline.gov), and the Guidelines International Network (www.g-i-n.net).

In January 2006, the ADAPTE working group merged with the Practice Guideline Evaluation and Adaptation Cycle (PGEAC) group, which had previously designed a framework for guideline adaptation that was pilot tested in an Ontarian project on community care of leg ulcers (7, 8). The framework has since been used by national (9–11) and local (12) groups. The group formed from the merger was renamed the ADAPTE Collaboration. Its goal was to support and evaluate a systematic approach for guideline adaptation. On the basis of the literature review, their collective experience, and extensive discussions during four 2-day meetings, they produced a manual (the ADAPTE manual) and a resource toolkit for guideline adaptation. In 2009, these resources were transferred to the Guidelines International Network (www.g-i-n.net). These materials were also used to address our questions.

The adaptation process is based on the following core principles:

- Use of widely accepted evidence-based principles for guideline development (13, 14)
- Development of reliable methods to ensure quality and validity of adapted guideline (15)
- Use of a participative approach involving all key stakeholders to foster acceptance and ownership (16)
- Explicit consideration of new users' context to ensure relevance for practice (17)
- Transparent reporting to promote confidence in the recommendations of the adapted guideline (15, 18)
- Use of a flexible format to accommodate specific needs and circumstances (19, 20)

## RESULTS

The literature review by the ADAPTE Collaboration identified 19 publications that reported models, practical examples, and/or experiences of guideline adaptation (3). Three types of guideline adaptation were distinguished as follows:

1. Adaptation as an alternative to *de novo* development (four publications)
2. Adaptation as part of an implementation process for an international guideline or a guideline developed for one country but appropriate for another country, taking into consideration contextual differences ("transcontextual adaptation") (eight publications)
3. Local adaptation of a national guideline, taking into consideration local context (seven publications)

The ADAPTE manual presents a generic approach that can be used as an alternative to *de novo* guideline development, as well as for customizing existing guidelines to suit the specific national or local context. The process described in the manual consists of 3 main phases, 9 modules, and 24 steps (Table 1).

### Before Getting Started

Suppose an association of health care professionals with scarce resources is interested in guideline adaptation because it could potentially save effort and time, while ensuring the availability of evidence reviews. Before getting started with the adaptation process, a series of questions should be considered (Table 2). First, what does the organization aim to achieve by engaging in guideline adaptation? Second, what is the specific health care issue that needs to be addressed by a guideline (e.g., is the issue the provider's patient care content, a system issue that hinders effective and efficient delivery of care, or a barrier to reimbursement for needed services?). Finally, is a guideline the right tool to achieve the goal? If the goal is education and training, guidelines can be useful. If the goal is to change provider or patient behavior or to improve quality of care, other interventions should be considered as well, such as organizational interventions (improving logistics and a multidisciplinary approach), financial (dis)incentives, legal measures (e.g., public smoking ban), or patient-mediated interventions. More details about effective implementation strategies are provided in the article by Grimshaw and colleagues in this series (21).

### 1. What Guidelines on COPD Are Available?

To explore whether adaptation is feasible, existing guidelines on the specific health topic covering the predefined key issues must be identified. Relevant characteristics of the retrieved guidelines are the developing organization and authors, date of publication, country and language of publication, and dates of the search used by the source guideline developers. Sources to search for existing guidelines include both print publications and websites, such as guideline clearinghouses, the websites of known guideline developers, and electronic databases (Table 3).

If a large number of potentially relevant guidelines is found, the following selection criteria may be considered:

- Selecting only evidence-based guidelines with a high-quality and transparent review process (e.g., those including a report on systematic literature searches and explicit explanatory links between individual recommendations and their supporting evidence)
- Selecting only guidelines produced by national and international agencies, assuming that they fulfill the first criterion
- Selecting only recently published or updated guidelines (e.g., published less than 2 years ago)

Many COPD guidelines are available, including several developed by prominent guideline organizations. Table 4 presents 10 publicly available guidelines published between 2004 and 2007: 2 international guidelines, 1 Latin American guideline, and 7 national guidelines. Seven of the guidelines are in English, one is in Spanish, one is in Dutch, and one is in German. All of these guidelines declare to have been produced according to the principles of evidence-based guideline development. Three of the guidelines are an update of earlier versions. The length of the guidelines varies from 11 to 267 pages and the number of references ranges from 92 to 560. A more detailed analysis of the quality of the guidelines can be performed at a later phase of the guideline adaptation process, when the guidelines that address the clinical questions formulated by the guideline working group are identified for further consideration.

### 2. How Should Guidelines Be Adapted?

The adaptation process starts with the formulation of specific questions that need to be addressed by the guideline. The number of questions may vary between 1 and 100 and will determine the workload. A large number of questions usually result from specifically targeted questions such as which medication should be added in a specific clinical situation. Broader questions are often more

**TABLE 1. PHASES, MODULES, AND STEPS IN ADAPTATION ACCORDING TO ADAPTE MANUAL**

| Phase | Modules | Steps |
|---|---|---|
| I. Setup | Preparation | • Establish an organizing committee<br>• Select a guideline topic<br>• Check whether adaptation is feasible<br>• Identify necessary resources and skills<br>• Complete tasks for the set-up phase<br>• Write adaptation plan |
| II. Adaptation | Scope and purpose | • Determine the health questions |
| | Search and screen | • Search for guidelines and other relevant documents<br>• Screen retrieved guidelines<br>• Reduce a large number of retrieved guidelines |
| | Assessment | • Assess guideline quality<br>• Assess guideline currency<br>• Assess guideline content<br>• Assess guideline consistency<br>• Assess acceptability/applicability of the recommendations |
| | Decision and selection | • Review assessments<br>• Select between guidelines and recommendations to create an adapted guideline |
| | Customization | • Prepare draft adapted guideline |
| III. Finalization | External review and acknowledgment | • External review by target users<br>• Consult with relevant endorsement bodies<br>• Consult with developers of source guidelines<br>• Acknowledge source documents |
| | Aftercare planning | • Plan scheduled review and update of adapted guideline |
| | Final production | • Produce final guidance document |

useful and more easily addressed (e.g., what is the usual hierarchy of COPD medications across stages?), but they may be less useful to those looking for specific answers. The quality of the preselected guidelines can be assessed with tools such as the Appraisal of Guidelines Research and Evaluation (AGREE) instrument (15). Only those guidelines with acceptable quality should be chosen for further analysis of the clinical content and the supporting evidence (*see below* for a discussion of how to determine acceptable quality).

***Which questions should be answered in the specific context?*** Once a broad topic area is identified, it is important to clarify the specific purpose and parameters of the chosen guideline topic by developing a series of structured key questions (22). Clear and focused key clinical questions (23) are necessary to successfully complete the adaptation process and will ensure that the final adapted guideline is applicable to the user's purpose in the user's context.

The use of the following four items (PICO) will help to define the clinical questions (5):

- Population concerned and characteristics of disease
- Intervention(s) (or diagnostic test) of interest
- Comparison of interest
- Outcomes of interest, including patient outcomes (e.g., survival or quality of life) and system outcomes (e.g., practice homogeneity)

**TABLE 2. CHECKLIST FOR (POTENTIAL) ADAPTERS**

Why have you chosen to focus on COPD? Please define the key issues about COPD in your country. For instance:
- Incidence of COPD is rising due to smoking or biomass (consider diversity issues as age, sex, SES)
- COPD is underdiagnosed and patients as well as (primary care) doctors are unaware of it
- Patients are over- or underweight or have comorbidity
- There is no access to affordable health care services
- Spirometry is available only in hospitals
- Inhalers/spacers are not available or patients do not apply it correctly
- Doctors are prescribing too many drugs and/or too expensive drugs
- Patients are undertreated
- Diagnosis and treatment of exacerbations are inappropriate and self-management is not promoted
- Adequate follow-up and a multidisciplinary approach are lacking
- Facilities for pulmonary rehabilitation are lacking or not used optimally
- Facilities for oxygen therapy are lacking or not used optimally
- There is insufficient end-of-life and palliative care

What are you attempting to achieve? What goal have you set?
Are guidelines the right tool to achieve your goal? Please also consider other interventions to meet your goal
Is there a (recent) guideline available that may fit in your context? If there are many, how would you select the best ones?
Who will be the target users of your guidelines? What are their needs and preferences?
Are you considering:
- Adoption of a guideline without any change, translation, restyling?
- Adaptation (with or without modifying the recommendations)?
- Using only the evidence summaries from existing guidelines (or existing systematic reviews)?

Who is responsible for the guideline development and maintenance?
What resources are available for development, dissemination, implementation, and updating?

*Definition of abbreviations*: COPD = chronic obstructive pulmonary disease; SES = socioeconomic status.

**TABLE 3. SELECTED GUIDELINE CLEARINGHOUSES AND OTHER SOURCES FOR GUIDELINES**

| Guideline Internet Site | URL |
|---|---|
| National Guideline Clearinghouse (NGC) | www.guideline.gov |
| Guidelines International Network (G-I-N) | www.g-i-n.net |
| Ontario Guidelines Advisory Committee (GAC) Recommended Clinical Practice Guidelines | www.gacguidelines.ca |
| Institute for Clinical Systems Improvement (ICSI) | www.icsi.org |
| National Institute for Health and Clinical Excellence (NICE) | www.nice.org.uk |
| New Zealand Guidelines Group (NZGG) | www.nzgg.org.nz |
| Scottish Intercollegiate Guidelines Network (SIGN) | www.sign.ac.uk |
| Haute Autorité de Santé (HAS) | www.has-sante.fr |
| German Agency for Quality in Medicine (AEZQ) | www.aezq.de |

In the context of guideline adaptation, the professionals to whom the guideline is targeted and the health care setting in which the guideline is to be implemented should also be considered.

***How should the quality of the guidelines be assessed?*** The AGREE instrument provides a framework for assessing the quality of clinical practice guidelines (15), but does not assess the clinical content of the recommendations. The instructions in the introduction of the instrument should be read carefully before starting the appraisal. Depending on the guideline, the appraisal process might take approximately 1.5 hours per guideline. Each guideline should be appraised by at least two, and preferably four, appraisers. Large differences in the scores of the same dimension across different guidelines identify the need for a specific discussion point.

Although the AGREE instrument does not provide thresholds for acceptable or unacceptable guidelines on the basis of quality, a comparison of rigor scores across guidelines can provide the panel with information to guide the selection of guidelines that should be included in the adaptation process. For example, the panel could decide on a cutoff point or rank the guidelines, once they see how the guidelines score on rigor (e.g., they may decide that any guideline scoring above 50% on the rigor dimension will be retained). Other options might be to keep all guidelines that score above the median score or all that score above the 60th percentile. Figure 1 shows an example of how quality scores of different guidelines could be compared with support guideline selection for the adaptation process. A limitation of an AGREE appraisal is that it relies on the report of the guideline development process. High scores do not exclude the possibility that research data may have been misinterpreted or that financial conflicts were not managed correctly. Further content analysis is needed as described below.

***How can the clinical content and validity of the recommendations be assessed?*** Once the guidelines of acceptable quality are selected, the clinical content and validity of the clinical recommendations must be assessed. Clinical expertise and familiarity with the empirical evidence in the specific disease area are needed for a reliable and valid assessment.

The guideline content can be presented by using recommendations matrices (23), grouped by the specific area covered. Quality scores on

**TABLE 4. OVERVIEW OF CLINICAL PRACTICE GUIDELINES ON CHRONIC OBSTRUCTIVE PULMONARY DISEASE PUBLISHED FROM 2004**

| Country | Organization | Type of Organization | Title of Guideline | Number of Pages | Number of References | Year of Publication |
|---|---|---|---|---|---|---|
| International | Global Initiative for Chronic Obstructive Lung Disease (GOLD) | International organization | Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease*† | 88 | 560 | 2006 (updated from 2001) |
| | American Thoracic Society and European Respiratory Society | Medical specialty societies | Standards for the diagnosis and treatment of patients with COPD§ | 222 | Approximately 400‡ | 2004 (updated from 1995) |
| Australia/ New Zealand | Australian Lung Foundation and the Thoracic Society of Australia and New Zealand | Public nonprofit organization and medical specialty society | The COPD-X Plan: Australian and New Zealand guidelines for the management of chronic obstructive pulmonary disease 2006† | 66 | 243 | April 2006 |
| Canada | Canadian Thoracic Society (CTS) | Medical specialty society | State of the Art Compendium: CTS recommendations for the management of chronic obstructive pulmonary disease§ | 59 | Approximately 400‡ | July 2004 |
| England/Wales | National Collaborating Centre for Chronic Conditions (NICE) | National government agency | National clinical guideline on management of chronic obstructive pulmonary disease in adults in primary and secondary care*† | 232 | 491 | February 2004 |
| Germany | German Agency for Quality in Medicine (AEZQ) | Private nonprofit organization | COPD—national disease management guideline (German language)† | 100 | 349 | December 2006 |
| The Netherlands | Dutch Institute for Healthcare Improvement (CBO) | Private nonprofit organization | COPD disease management (Dutch language)† | 267 | Approximately 450‡ | July 2005 |
| Singapore | Singapore Ministry of Health | National government agency | Chronic obstructive pulmonary disease*† | 84 | 155 | October 2006 |
| United States | Institute for Clinical Systems Improvement | Private nonprofit organization | Chronic obstructive pulmonary disease*† | 65 | 125 | January 2007 (updated from 2001) |
| Latin America | Latin American Thoracic Society | Medical specialty society | Chronic obstructive pulmonary disease: infectious exacerbation | 11 | 92 | 2004 |

* Selected from U.S. National Guideline Clearinghouse (www.guideline.gov), MESH Category "chronic obstructive pulmonary disease."

† Selected from Guidelines International network (www.g-i-n.net), Health topics collection MESH Category "chronic obstructive pulmonary disease" (C08.381.495.389).

‡ References listed at the end of each chapter, so there may be considerable overlap.

§ Selected from www.google.com, "COPD guidelines" (first 10 hits).

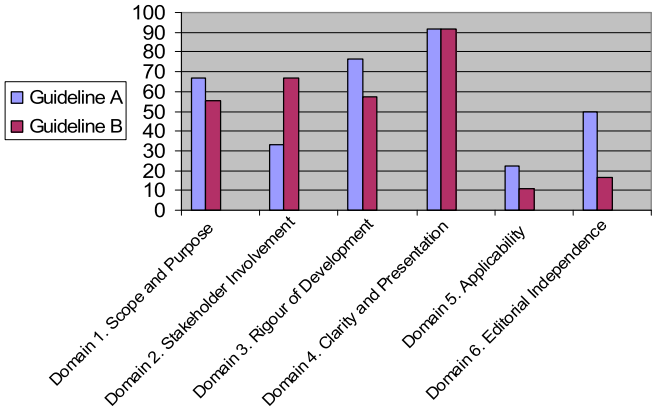


*Figure 1.* Example of Appraisal of Guidelines Research and Evaluation (AGREE) domain scores of two guidelines on chronic obstructive pulmonary disease.

the AGREE instrument, supporting studies, and levels of evidence could be added. This facilitates comparison of the recommendations and can support decision-making. Examples of recommendations matrices on COPD are available at the U.S. National Guideline Clearinghouse website (www.guideline.gov; click on guideline syntheses, and then select COPD). Lacasse and colleagues did similar work in 2001 (24).

Assessment of the validity of the recommendations includes three evaluations:

- Consistency between the search strategy and the selection of evidence supporting the recommendations
- Consistency between the selected evidence and how developers summarize and interpret this evidence
- Consistency between the interpretation of the evidence and the recommendations

The resource toolkit produced along with the ADAPTE manual includes several criteria for each of these evaluations. These may require the gathering of original evidence supporting the interpretations and recommendations in the guideline. The GRADE approach can be used to assess the consistency between the evidence and the recommendations (25). It is possible to reclassify the levels of evidence across multiple source guidelines if sufficient information on the original research studies is available, including considered judgment. The GRADE approach is discussed in more detail in two other articles of this series (26, 27).

***How should the acceptability, applicability, and transferability of the recommendations be assessed?*** The acceptability and applicability of recommendations from a guideline depend on the organizational and cultural context in which the recommendations are to be used. This includes the availability of health services, expertise, equipment, techniques, resources, and organization of health services, as well as patient population characteristics, cultural beliefs, and value judgments. For example, if certain diagnostic or therapeutic interventions recommended in the original guidelines are not available, the working group will need to decide whether there are acceptable (evidence-based) alternatives. If no alternatives are available, strategies to encourage the availability of the recommended services must be developed. Strategies may include political pressure to increase funding or decrease costs. For example, COPD diagnosis and management require spirometry, which is unavailable in primary care in many countries. An evidence-based guideline supported by all relevant stakeholders may help to speed the introduction of spirometry in these health care settings.

***How should final decisions be made about the adoption and adaptation of recommendations?*** A consensus procedure is needed to decide which guidelines and which recommendations will be selected for adoption or modification. This may be an eclectic process in which aspects from various guidelines may be selected for the adapted guideline, based on the quality of evidence and relevance of the recommendation in the source guidelines. New evidence reported after publication of the guidelines should be considered to determine whether any of the recommendations should be updated. Any modifications to the recommendations from the source documents should be carefully documented and the evidence supporting the modification provided, along with the quality and references.

### 3. How Should the Use of Guidelines Be Evaluated?

There is an increasing tendency to hold health care professionals, health care organizations, and hospitals accountable for the quality of care that they provide through public reporting of adherence to performance measures, as well as financial incentives and disincentives. Guideline recommendations are often used as the basis for performance measures, and this can offer an opportunity to monitor the use of guidelines in clinical practice. Programs that attempt to improve quality of care through the use of performance measures may produce important improvements in patient outcomes. However, there is increasing recognition of unintended consequences of such programs. For example, efforts to increase adherence to recommendations for timely antibiotics and hospital-based vaccination for pneumonia may have resulted in the overuse of antibiotics (28) and vaccinations (29). There is also concern regarding diminished access to care for socioeconomically disadvantaged or high-risk patients who are at risk for poor outcomes that may be "blamed" on the physician caring for such patients (30). The potential for such negative unintended consequences demands that programs designed to increase adherence to guideline recommendations be based on the highest quality evidence and that adhering to the recommendation improves patient outcomes after considering both benefits and downsides or at the very least improve the quality or process of care.

With respect to COPD, it is difficult to imagine that increased adherence to a recommendation for tobacco cessation could lead to negative consequences that outweigh the tremendous benefits. However, another proposed performance measure, regular performance of spirometry in patients diagnosed with mild or moderate COPD, may not pass that test. This recommendation has not been shown to improve patient outcomes, and although it is unlikely to directly harm patients, it may increase health care costs if spirometry is regularly performed in patients with stable mild to moderate COPD who have already discontinued cigarette smoking. Therefore, guideline authors should be aware of potential limitations and unintended consequences. Guideline authors could reduce adverse effects by providing a set of performance measures derived from the strong recommendations within the guideline.

### 4. How Should Guidelines Be Efficiently Kept Up-to-Date?

Guideline development should be considered an ongoing process (31). Shekelle and colleagues recommended reassessing guidelines at least every 3 years (32). As COPD is a disease area that attracts many innovations, more frequent guideline updating may be necessary and the approach of "living guidelines" should be adopted. This could include a yearly update of the guideline by a fixed panel, with or without rotating membership.

Performance evaluations and feedback on guideline use in clinical practice may be used in updating the guideline. Information and feedback from guideline users may improve the acceptance and implementability of the recommendations in subsequent versions of the guideline. Similar to developing a new guideline, recent existing guidelines can be used in updating adapted guidelines. Contacting other guideline developers who are active in the same disease area and exchanging evidence summaries could facilitate the process of updating and prevent duplication of effort.

## DISCUSSION

Adaptation of guidelines in the field of COPD is an attractive approach as an alternative for *de novo* guideline development, as many high-quality guidelines on the topic are widely available. Two international guidelines, the GOLD guideline and the ATS/ERS guideline, were particularly well produced for this specific purpose. The GOLD guideline, for instance, is available in several translations, facilitating its use in different countries.

Guideline adaptation, however, not only includes translation, but also considers the context, relevant stakeholders, potential guideline users, and cultural values. In-depth analysis of the guidelines is needed to assess the applicability and transferability of the recommendations into one's own context. For example, the GOLD guideline has been translated into Polish and the recommendations adapted for local needs. At the same time, such adaptations bear the risk that they may become disconnected from the original recommendations.

Whenever guideline adaptation is considered, a systematic approach should be followed to ensure high-quality outcomes. The systematic approach requires resources and skilled participants. No evidence is available on the extent to which guideline adaptation may save time or resources. Until evidence is available, pragmatic decisions may be made to reduce the workload, provided that the process is reported transparently. For instance, there are groups that have limited resources and no currently adapted guidelines. They may be able to skip some of the adaptation processes and concentrate on customization if other groups have already gone through all of the adaptation steps and have published their findings. Centralized resources that allow adaptation could reduce the time and effort for those with limited resources.

The adaptation approach described in this article has been evaluated among 144 registrants of the ADAPTE website (4). The ADAPTE process and manual were rated as clear and comprehensive by a majority of respondents. Most (89%) respondents were expecting benefit from using the ADAPTE process; 75% of participants were expecting the quality of the customized guidelines to be increased; and less than 30% were expecting the timeframe and resource use to be decreased in comparison with their current procedures. Several comments indicated that participants considered ADAPTE to be more rigorous than their current processes. The most often reported anticipated barriers to using the adaptation process were low-quality source guidelines, lack of appropriate source guidelines, and difficulties adapting source guidelines to their own context of use. Additional reported barriers included lack of expertise and time, as well as limited instructions for guideline implementation in the ADAPTE manual. It can be concluded that, as with other guideline methods, users should take into account that the use of ADAPTE might require a learning process, in particular for those not familiar with structured guideline methods.

In addition to the usefulness of helping guideline developers produce and update guidelines, guideline adaptation provides an opportunity for national and international collaborations among organizations that share common issues and an opportunity to investigate more efficient ways to provide guidelines to multiple countries and populations (1). The approach is also consistent with the aims of the World Health Organization, which provides local support for adapting and implementing recommendations by developing tools, building capacity, learning from international experience, and through international networks that support evidence-informed health policies (2).

***Author Disclosures***: J.S.B. received lecture fees from Sanofi-Aventis. A.A. consulted with Bayer-Schering Pharma, Boehringer Ingelheim, Dey Pharma, Forest, GlaxoSmithKline, Pfizer, and Schering-Plough, and served on advisory committees of Boehringer Ingelheim, Dey, and GlaxoSmithKline. He received lecture fees from AstraZeneca, Bayer-Schering Pharma, Pfizer, and Schering-Plough, and research grants from Eli Lilly, GlaxoSmithKline, and Pneuma. He is a member of the ATS/ERS COPD Guideline Development Committee and a member of GOLD Executive and Scientific committees. P.N.B. reported no commercial interests relevant to subject matter. A. A.C. was employed by the World Health Organization. B.F. and I.D.G. were members of the ADAPTE Collaboration. M.M. consulted with Qualidigm and Pfizer, was on advisory committees of Pfizer, and received lecture fees from Pfizer. M.W. consulted with GlaxoSmithKline and received lecture fees from Pfizer. B.P.Y. consulted with Boehringer Ingelheim, GlaxoSmithKline, Novartis and Pfizer, and was on advisory committees of Astra Zeneca, Merck, Pfizer and Schering. She received grants from Astra Zeneca, Boehringer Ingelheim, Merck, Novartis and Pfizer. All workshop participants disclosed receipt of honoraria and travel reimbursement from the American Thoracic Society.

### References

1. Schünemann HJ, Fretheim A, Oxman AD. Improving the use of research evidence in guideline development. 13. Applicability, transferability and adaptation. *Health Res Policy Syst* 2006;4:25.
2. Ollenschläger G, Marshall C, Qureshi S, Rosenbrand K, Burgers J, Makela M, Slutsky J. Improving the quality of health care: using international collaboration to inform guideline programmes by founding the Guidelines International Network (G-I-N). *Qual Saf Health Care* 2004;13:455–460.
3. Harrison MB, Légaré F, Graham ID, Fervers B. Adapting clinical practice guidelines to local context and assessing barriers to their use. *CMAJ* 2010;182:E78–E84.
4. Fervers B, Burgers JS, Voellinger R, Brouwers M, Browman GP, Graham ID, Harrison MB, Latreille J, Mlika-Cabane N, Paquet L, Zitzelsberger L, Burnand B; ADAPTE Collaboration. Guideline adaptation: an approach to enhance efficiency in guideline development and improve utilisation. *BMJ Qual Saf* 2011;20:228–236.
5. Fervers B, Burgers JS, Haugh M, Latreille J, Mlika-Cabanne N, Paquet L, Coulombe M, Poirier M, Burnand B. Adaptation of clinical guidelines: literature review and proposition for a framework and procedure. *Int J Qual Health Care* 2006;18:167–176.
6. Schunemann HJ, Woodhead M, Anzueto A, Buist S, Macnee W, Rabe KF, Heffner J. A vision statement on guideline development for respiratory disease: the example of COPD. *Lancet* 2009;373:774–779.
7. Graham ID, Harrison MB, Lorimer K, Piercianowski T, Friedberg E, Buchanan M, Harris C. Adapting national and international leg ulcer practice guidelines for local use: the Ontario leg ulcer community care protocol. *Adv Skin Wound Care* 2005a;18:307–318.
8. Harrison MB, Graham ID, Lorimer K, Friedberg E, Pierschianowski T, Brandys T. Leg-ulcer care in the community, before and after implementation of an evidence-based service. *CMAJ* 2005;172: 1447–1452.
9. Canadian Stroke Network; Heart & Stroke Foundation of Canada. The Canadian stroke strategy, 2012: Canadian best practice recommendations for stroke care [Internet] [accessed September 2012]. Available from: http://www.strokebestpractices.ca/
10. Canadian Stroke Network. Stroke Canada optimization of rehabilitation through evidence (SCORE). Score evidence based recommendations for the upper and lower extremities and risk assessment post stroke, 2005 [Internet] [accessed September 2012]. Toronto: Canadian Stroke Network-Réseau Canadien contre les Accidents Cerebrovasculaires. Available from: http://www.canadianstrokenetwork.ca/index.php/tools/
11. Elit LM, Johnston M, Brouwers M, Fung-Kee-Fung M, Browman G, Graham ID; on behalf of the Society of Gynecologic Oncologists of Canada and the Canadian Strategy for Cancer Control Clinical Practice Guidelines Action Group. Promoting best gynecologic oncology practice: a role for the Society of Gynecologic Oncologists of Canada. *Curr Oncol* 2006;13:1–5.

12. Sprague AE, Oppenheimer L, McCabe L, Brownlee J, Graham ID, Davies B. The Ottawa Hospital's clinical practice guideline for the second stage of labour. *J Obstet Gynaecol Can* 2006;28:769–779.
13. Sackett D. Evidence-based medicine: how to practice and teach EBM, 2nd ed. Philadelphia: Churchill Livingstone; 2000.
14. Scottish Intercollegiate Guidelines Network. SIGN 50: a guideline developer's handbook, 2008 [Internet] [accessed September 2012]. Available from: http://www.sign.ac.uk/pdf/sign50.pdf
15. AGREE Collaboration. Development and validation of an international appraisal instrument for assessing the quality of clinical practice guidelines: the AGREE project. *Qual Saf Health Care* 2003;12:18–23.
16. National Institute for Health and Clinical Excellence (NICE). How NICE clinical guidelines are developed: an overview for stakeholders, the public and the NHS, 4th ed. London: NICE; 2009.
17. Verkerk K, Van Veenendaal H, Severens JL, Hendriks EJ, Burgers JS. Considered judgement in evidence-based guideline development. *Int J Qual Health Care* 2006;18:365–369.
18. Shiffman RN, Shekelle P, Overhage JM, Slutsky J, Grimshaw J, Deshpande AM. Standardized reporting of clinical practice guidelines: a proposal from the Conference on Guideline Standardization. *Ann Intern Med* 2003;139:493–498.
19. Grol R, Dalhuijsen J, Thomas S, Veld C, Rutten G, Mokkink H. Attributes of clinical guidelines that influence use of guidelines in general practice: observational study. *BMJ* 1998;317:858–861.
20. Burgers JS, Grol RP, Zaat JO, Spies TH, van der Bij AK, Mokkink HG. Characteristics of effective clinical guidelines for general practice. *Br J Gen Pract* 2003;53:15–19.
21. Grimshaw G, Schünemann HJ, Burgers JS, Cruz AA, Heffner J, Metersky M, Cook D. Disseminating and implementing guidelines. Article 13 of Integrating and coordinating efforts in COPD guideline development. An Official ATS/ERS Workshop Report. *Proc Am Thorac Soc* 2012;9:298–303.
22. Graham ID, Harrison MB. Evaluation and adaptation of clinical practice guidelines. *Evid Based Nurs* 2005;8:68–72.
23. Graham ID, Harrison MB, Brouwers M. Evaluating and adapting practice guidelines for local use: a conceptual framework. In: Pickering S, Thompson J, editors. Clinical governance in practice. London: Harcourt Publishers; 2003. pp. 213–229.
24. Lacasse Y, Ferreira I, Brooks D, Newman T, Goldstein RS. Critical appraisal of clinical practice guidelines targeting chronic obstructive pulmonary disease. *Arch Intern Med* 2001;161:69–74.
25. GRADE Working Group. Grading quality of evidence and strength of recommendations. *BMJ* 2004;328:1490–1497.
26. Guyatt G, Akl E, Oxman A, Wilson K, Puhan MA, Wilt T, Gutterman D, Woodhead M, Antman E, Schünemann HJ. 7. Synthesis, grading and presentation of evidence in guidelines. Article 7 of Integrating and coordinating efforts in COPD guideline development. An Official ATS/ERS Workshop Report. *Proc Am Thorac Soc* 2012;9: 256–261.
27. Schunemann H, Oxman A, Akl EA, Brozek J, Montori V, Heffner J, Hill S, Woodhead M, Campos-Outcalt D, Alderson P, *et al*. Moving from evidence to developing recommendation in guidelines. Article 11 of Integrating and coordinating efforts in COPD guideline development. An Official ATS/ERS Workshop Report. *Proc Am Thorac Soc* 2012;9:282–292.
28. Kanwar M, Brar N, Khatib R, Fakih MG. Misdiagnosis of community-acquired pneumonia and inappropriate utilization of antibiotics: side effects of the 4-h antibiotic administration rule. *Chest* 2007;6:1865–1869.
29. Metersky M, Ho SY, Petrillo MK. Does hospital-based influenza vaccination result in unintended revaccination? *Am J Med* 2004;116:854–855.
30. Epstein AM, Lee TH, Hamel MB. Paying physicians for high-quality care. *N Engl J Med* 2004;350:406–410.
31. Oxman AD, Schunemann HJ, Fretheim A. Improving the use of research evidence in guideline development. 16. Evaluation. *Health Res Policy Syst* 2006;4:28.
32. Shekelle P, Eccles MP, Grimshaw JM, Woolf SH. When should clinical guidelines be updated? *BMJ* 2001;323:155–157.