

## Evidence-Based Medicine



## An Introduction to an Evidence-Based Approach to Interventional Techniques in the Management of Chronic Spinal Pain

Laxmaiah Manchikanti, MD<sup>1</sup>, Vijay Singh, MD<sup>2</sup>, Standiford Helm II, MD<sup>3</sup>,  
David M. Schultz, MD<sup>4</sup>, Sukdeb Datta, MD<sup>5</sup>, and Joshua Hirsch, MD<sup>6</sup>

From: <sup>1</sup>Pain Management Center of Paducah, Paducah, KY; <sup>2</sup>Pain Diagnostics Associates, Niagara, WI; <sup>3</sup>Pacific Coast Pain Management Center, Laguna Hills, CA; <sup>4</sup>Medical Advanced Pain Specialists, Minneapolis, MN; <sup>5</sup>Vanderbilt University Medical Center, Nashville, TN; and <sup>6</sup>Massachusetts General Hospital and Harvard Medical School, Boston, MA

Dr. Manchikanti is Medical Director of the Pain Management Center of Paducah, Paducah, KY. Dr. Singh is Medical Director of Pain Diagnostics Associates, Niagara, WI. Dr. Helm is a Medical Director, Pacific Coast Pain Management Center, Laguna Hills, CA. Dr. Schultz is the Medical Director of Medical Advanced Pain Specialists, Minneapolis, MN. Dr. Datta is Director, Vanderbilt University Interventional Pain Program, Associate Professor, Dept. of Anesthesiology, Vanderbilt University Medical Center, Nashville, TN. Dr. Hirsch is Chief of Minimally Invasive Spine Surgery, Depts. of Radiology and Neurosurgery, Massachusetts General Hospital and Assistant Professor of Radiology, Harvard Medical School, Boston, MA.

Address correspondence:  
Laxmaiah Manchikanti, MD  
2831 Lone Oak Road  
Paducah, Kentucky 42003  
E-mail: drlm@thepainmd.com

Disclaimer: There was no external funding in the preparation of this manuscript.

No statement on this article should be construed as an official position of ASIPP. Dr. Hirsch is a consultant for Cardinal Healthcare and Medtronic, he has previously served as a consultant for Arthrocare, he serves on the Steering Committee for KAVIAR trial (volunteer position), and on the Data and Safety Monitoring Board (DSMB): CEEP trial (volunteer position). Dr. Schultz receives funding from Medtronic. Conflict of interest: None.

Manuscript received: 04/07/2009  
Accepted for publication: 05/14/2009

Free full manuscript:  
[www.painphysicianjournal.com](http://www.painphysicianjournal.com)

Practice guidelines are systematically developed statements to assist practitioners and patients in making decisions about appropriate health care for specific clinical circumstances. Clinical practice guidelines present statements of best practice based on a thorough evaluation of the evidence from published studies on the outcomes of treatment. In November 1989, Congress mandated the creation of the Agency for Healthcare Policy and Research (AHCPR). AHCPR was given broad responsibility for supporting research, data development, and related activities. Associated with this mandate, the National Academy of Sciences published a document indicating that guidelines are expected to enhance the quality, appropriateness, and effectiveness of health care services.

Guidelines as a whole have been characterized by multiple conflicts in terminology and technique. These conflicts are notable for the confusion they create and for what they reflect about differences in values, experiences, and interest among different parties. Despite this confusion, public and private development of guidelines is growing exponentially. There are only limited means to coordinate these guidelines in order to resolve inconsistencies, fill in gaps, track applications and results, and assess the soundness of individual guidelines. Significant diversity exists in clinical practice guidelines. The inconsistency amongst guidelines arises from variations in values, tolerance for risks, preferences, expertise, and conflicts of interest.

In 2000, the American Society of Interventional Pain Physicians (ASIPP) first created treatment guidelines to help practitioners. There have been 4 subsequent updates. These guidelines address the issues of systematic evaluation and ongoing care of chronic or persistent pain, and provide information about the scientific basis of recommended procedures. These guidelines are expected to increase patient compliance, dispel misconceptions among providers and patients, manage patient expectations reasonably, and form the basis of a therapeutic partnership between the patient, the provider, and payors.

The ASIPP guidelines are based on evidence-based medicine (EBM). EBM is in turn based on 4 basic contingencies: the recognition of the patient's problem and the construction of a structured clinical question; the ability to efficiently and effectively search the medical literature to retrieve the best available evidence to answer the clinical question; clinical appraisal of the evidence; and integration of the evidence with all aspects of the individual patient's decision-making to determine the best clinical care of the patient. Evidence synthesis for guidelines includes the review of all relevant systematic reviews and individual articles, grading them for relevance, methodologic quality, consistency, and recommendations.

**Key words:** Evidence-based medicine, clinical practice guidelines, critical appraisal, guideline development, interventional pain management, interventional techniques, evidence synthesis, clinical relevance, grading recommendations, systematic reviews

**Pain Physician 2009; 12:E1-E33**

Clinical guidelines are a constructive response to the reality that practicing physicians require assistance into assimilating and applying the exponentially expanding, often contradictory, body of medical knowledge (1). Clinical practice guidelines attempt to define practices that meet the needs of most patients under most circumstances. They do not attempt to supplant the independent judgment of clinicians in responding to particular clinical situations (2). Ideally, the specific clinical recommendations that are contained within the practice guidelines have been systematically developed by panels of experts who have access to the available evidence, have an understanding of the clinical problem, and have clinical experience with the subject procedure and the relevant research methods to make considered judgments. These panels are expected to be objective and to produce recommendations that are unbiased, up-to-date, and free from conflict of interest. Consequently, guidelines are widely perceived as evidence based, rather than authority based, and therefore unbiased and valid (1). When strict adherence to the formal process and principles of evidence synthesis and grading and guideline development by a specialty society relevant to the subject is performed by experts in the subject, the guidelines thus produced are assumed to have the same level of certainty and security as conclusions generated by the conventional scientific method (1). For many clinicians, guidelines have become the final arbiters of care.

The Institute of Medicine (IOM) defines clinical guidelines as, "systematically developed statements to assist practitioner and patient decisions about appropriate health care for specific clinical circumstances" (3). In essence, guidelines enable the implementation of evidence-based medicine (EBM) during medical decision with the goal of encouraging effective care. EBM is commonly defined as, "the conscientious, explicit, and judicious use of current best evidence in making decisions about the care of individual patients" (4). The term is loosely used and can refer to anything from conducting a statistical meta-analysis of accumulated research, to promoting randomized clinical trials, to supporting uniform reporting styles for research, to a personal orientation toward critical self-evaluation (5). EBM was initially defined in opposition to clinical experience, while more recent definitions have emphasized the complementary character of the 2 and have aimed to improve clinical experience with better evidence (6).

Both clinical practice guidelines and EBM are experiencing exponential growth (7-9). Indeed, any group of individuals can designate itself as an EBM or guideline group. Different groups have reviewed the same procedure or problem in interventional pain management and reached vastly different conclusions (10-97). Consequently, it is clear that the process of preparation of EBM manuscripts and guidelines is insecure, and replication, which is the distinguishing characteristic of scientific knowledge and an essential test of the validity of any scientific statement, is basically impossible.

As a result, researchers, clinicians, professional organizations, and governments in the United States and other countries are looking for a sensible approach to health care with practicable and replicable EBM. However, each segment has their own interpretation and agenda, which is not based on science and best care for the patient, but seemingly on other factors. While the actual value of the evidence is related to the application and circumstances in which it will be used and for whom, such evidence may or may not have relevance. It is also essential to remember that the value of evidence is only as good as the type of evidence reviewed, methodology utilized, knowledge and experience of the reviewers, and many other factors, including bias, self-interest, and economics. EBM begins with the assertion that it is a shift in medical paradigms and is about solving clinical problems (98-100). A formal set of rules must complement medical training and common sense for clinicians to interpret the results of clinical research effectively (98-100). Thus, knowing the tools of evidence-based practice is necessary, but not sufficient, for delivering the highest quality of patient care. It therefore continues to be a challenge for EBM, and for interventional pain management, to better integrate the new science with the time-honored craft of caring for the sick (99). Even though some have characterized EBM as a stick by which policy-makers and academicians beat clinicians, there is an extensive role for EBM and clinical guidelines based on EBM in interventional pain management (101-105).

The influence of EBM and clinical guidelines on clinical practice and health policy is enormous. However, the process of guideline development has been, and remains, essentially unregulated (1). Sniderman and Furberg (1) have examined the sources of guideline authority; identified major limitations of the present process; addressed the issue of conflict of interest,

both for the individuals who staff the committees and the organizations that govern them; and provided suggestions for reform that may help improve the conduct of the process. However, strict implementation of these reforms may complicate EBM and guideline synthesis further with extreme views based on the sponsor of the guidelines. It has been stated that critics of EBM mostly come from within the medical professions (5). In addition to the many scientific problems of creating sound guidelines when evidence is weak, they stress the destructive effects of standards at the local level. Consequently, in an age of mandated cost control, managed care, and resource limitation, many practitioners “in the trenches” believe that instead of revolutionizing care, EBM threatens to bring about stagnation and bland uniformity, denial of coverage, derogatorily characterized as “cookbook medicine” (5). Ironically, EBM may also result in a lower standard of safety and economy by deskilling practitioners and increasing costs. In contrast, supporters tend to see EBM and guidelines as a panacea for the problems of rising costs, and the inequity and variability plaguing the health care field (5). The supporters contend that individual clinicians using EBM will be able to draw upon the objective experience of many researchers working with accepted scientific standards of evidence. Thus, this evidence is related to an assessment of the patient’s circumstances and the practitioner’s clinical experience, improving efficacy by allowing providers to filter scarce resources away from ineffective clinical practices and toward practices whose effectiveness has been conclusively shown.

The notion that EBM promises to create better informed patients and clinicians by offering collectively agreed-upon and publicly available information about treatment options is contradicted by a significant proportion of physician providers. In practice, EBM clinical practice guidelines are created by a small group of interested parties.

Appropriately developed guidelines must incorporate validity, reliability, reproducibility, clinical applicability and flexibility, clarity, development through a multidisciplinary process, scheduled reviews, and documentation (2,3,7,106). When appropriately applied, rigorously developed guidelines have the potential to reduce undesirable practice variation, reduce the use of services that are of minimal or questionable value, increase utilization of services that are effective but underused, and target services to those populations most likely to benefit (107-112).

## **1.0 DEFINITIONS**

---

### **1.1 Chronic Pain**

Chronic pain is defined as a complex and multifactorial phenomenon with pain that persists 6 months after an injury and/or beyond the usual course of an acute disease or a reasonable time for a comparable injury to heal, that is associated with chronic pathologic processes that cause continuous or intermittent pain for months or years, that may continue in the presence or absence of demonstrable pathology and may not be amenable to routine pain control methods with healing never occurring (113).

### **1.2 Interventional Pain Management**

The National Uniform Claims Committee (NUCC) (114) defined interventional pain management as the discipline of medicine devoted to the diagnosis and treatment of pain and related disorders by the application of interventional techniques in managing subacute, chronic, persistent, and intractable pain, independently or in conjunction with other modalities of treatments.

### **1.3 Interventional Techniques**

The Medicare Payment Advisory Commission (MedPAC) (115) described interventional techniques as minimally invasive procedures, such as needle placement of drugs in targeted areas, ablation of targeted nerves, and some surgical techniques, such as discectomy and the implantation of intrathecal infusion pumps and spinal cord stimulators.

### **1.4 Evidence-based Medicine**

EBM is defined as a conscientious, explicit, and judicious use of current best evidence in making decisions about care of individual patients (4).

### **1.5 Guidelines**

The IOM defined clinical guidelines as “systematically developed statements to assist practitioner and patient decisions about appropriate health care for specific clinical circumstances” (3).

### **1.6 Systematic Reviews and Meta-analyses**

A systematic review is defined as “the application of scientific strategies that limit bias by the systematic assembly, critical appraisal, and synthesis of all relevant studies on a specific topic” (116).

Meta-analysis, in contrast to a systematic review, is the statistical pooling of data across studies

to generate a summary (pooled estimates of effects) (117,118).

### **1.7 Randomized Controlled Trials**

A randomized controlled trial (RCT) is defined as any research study that randomly assigns human participants or groups of humans to one or more health-related interventions versus a placebo to evaluate the effects on health outcomes (119-122).

### **1.8 Observational Studies**

An observational study is defined as an etiologic or effectiveness study, a cross-sectional study, a case series, a case-control design, a design with historical controls, or a cohort design (123,124).

### **1.9 Diagnostic Studies**

A diagnostic study is a study of diagnostic accuracy, in which results from one or more tests are compared with the results obtained with the reference standard on the same subject (125).

## **2.0 BASIC CONSIDERATIONS**

---

### **2.1 Purpose**

Evidence-based clinical practice guidelines for interventional techniques in the management of chronic spinal pain are statements developed to improve the quality of care, patient access, treatment outcomes, appropriateness of care, efficiency and effectiveness, and achieve cost containment by improving the cost-benefit ratio (22-26).

### **2.2 Focus**

These updated and revised guidelines focus on a range of interventions that are the essential elements of effective management of chronic spinal pain.

### **2.3 Objectives**

The objectives of the American Society of Interventional Pain Physicians (ASIPP) guidelines for interventional techniques are to provide a set of recommendations that can support existing and future guidelines by:

1. Providing strategies to manage chronic spinal pain and/or its consequences in the general populations and in workers to improve the quality of clinical care.
2. Providing recommendations that are generally acceptable to a wide range of specialties and agencies.

3. Developing methods that are sound and transparent and highlighting the areas where further research is needed by noting deficiencies in knowledge.
4. Utilization of a process which is valid, reliable, reproducible, clinically applicable, and flexible, providing clarity with a multidisciplinary process with documentation of the process in developing guidelines, along with a scheduled review.
5. Systematically assessing the clinical and cost effectiveness of treatments and management strategies with an evidence-based approach through the use of systematic reviews, existing evidence-based guidelines, and individual clinical studies.
6. Increasing compliance, dispelling misconceptions, contributing to appropriate patient expectations, and facilitating the improved relationship between patients, physicians, and payors.

### **2.4 Population and Preferences**

The population covered by these guidelines includes all patients suffering with chronic spinal pain eligible to undergo commonly utilized and effective interventional technique(s). The treatment plan must be taken into consideration as well as the evidence, patient preferences, and risk-benefit ratio.

### **2.5 Implementation and Review**

The dates for implementation and review were established:

- ◆ Effective date – May 1, 2009
- ◆ Expiration date – April 31, 2012
- ◆ Scheduled review – April 1, 2011

### **2.6 Application**

While these guidelines may be applied by any specialty, they are specifically intended for use by interventional pain physicians. These guidelines do not constitute inflexible treatment recommendations. It is expected that a provider will establish a plan of care on a case-by-case basis, taking into account an individual patient's medical condition, personal needs, and preferences, and the physician's experience. Based on an individual patient's needs, treatment different from that outlined here could be warranted. Consequently, these guidelines do not represent a "standard of care."

The goal of these guidelines is to provide practitioners and payors information to determine whether the available evidence supports the notion of a "stan-

standard" for interventional techniques. "Standard" refers to what is applicable to the majority of patients, with a preference for patient convenience and ease of administration without compromising treatment efficacy or morbidity (126). It is essential to recognize the difference between "standard" and "standard of care," as utilized as a legal definition.

### **3.0 RATIONALE**

Despite advances in biomedical knowledge and the highest per capita health care expenditures in the world, the quality and outcomes of health care vary dramatically across the United States (7,127). The United States Government Accountability Office (GAO) in a letter to the Senate on September 26, 2008, informed the Senate that rapid spending growth per Medicare Part B – which covers physician and other patient services – has heightened concerns about the long-range fiscal sustainability of Medicare (128). Further, Medicare Part B expenditures are expected to increase over the next decade at an average annual rate of about 8%, which is faster than the projected 4.8% annual growth rate in the national economy over this time period (107,129). Accordingly, the trend to develop and implement research in support of evidence-based practice has been a focus of medical practice for the past decade. For example, in the modern era, the central premise is that decisions about the care of individual patients should be based on "the conscientious, explicit, and judicious use of current best evidence" (4). This means that individual clinical expertise should be integrated with the best information from scientifically based, systematic research, and should be applied in light of the patient's unique values and circumstances (130).

Towards these ends, ASIPP has provided evidence-based guidelines (22-26), methodology for evidence synthesis (27-33), systematic reviews (34-58), and critical analysis (60-63) based on a methodical critical appraisal of existing data using established and uniform criteria.

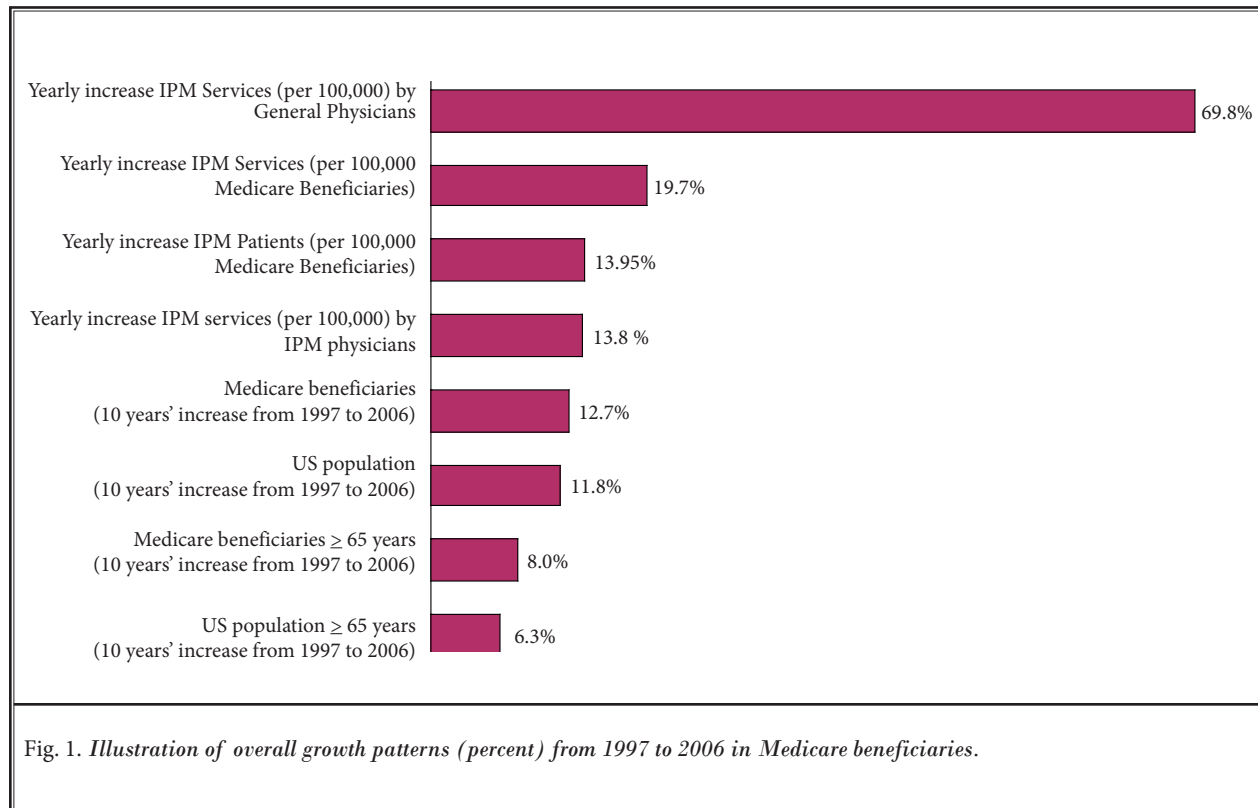
Innovations in health care are escalating at an astounding pace, adding complexity to the broad arena of health care interventions and systems (127,130-132). The demonstration of pervasive, persistent, and unexplained variability in clinical practice, high rates of inappropriate care, and escalating health care expenditures have fueled a steadily increasing demand for evidence of clinical effectiveness (4,7,64,98,99,107, 127,133-145). The demand expects a body of evidence

regarding safety, effectiveness, appropriate indications, cost-effectiveness, and other attributes of medical care. Numerous guidelines have been released with increasing regulations, both in the public and private sectors (2,7,10-26,60-63,65).

As an emerging speciality, interventional pain management faces multiple problems which may be disproportionate compared to established medical specialities. Interventional pain management is faced with increasing appropriate utilization of effective safe techniques due to its emergent nature as well as potentially inappropriate care that may be ineffective or unsafe (107,133-139). The available evidence at the present time documents a wide degree of variance in the definition and the practice of medicine in general and interventional pain management in particular (7,22-26,107,127,132-139). The application of interventional techniques by physicians of different specialties is highly variable for even the most commonly performed procedures and treated conditions (7,60-63,107,127,132-139,146,147).

The data of interventional techniques in the Medicare population from 1997 to 2006 shows an increase of 235% (107). However, the increase in expenditures in the Medicare population for IPM procedures has been even more significant with a 555% increase over the same period from \$113,087,880 in 1997 to \$740,719,400 in 2006. Yet, during the same period, the U.S. population increased by 12% and the Medicare population increased by 13% as a proportion of the population (107). The number of patients receiving interventional techniques increased by 169% from 1997 to 2006 (Fig. 1). This increase in utilization paralleled the advancement of new and innovative fluoroscopic injection techniques and the evolution of interventional pain management into a distinct medical specialty now designated with modifier -09 by the Center for Medicare Services (CMS).

The Office of Inspector General (OIG), Department of Health and Human Services, has reported startling data: Medicare Part B payments for facet joint injections have increased from \$141 million in 2003 to \$307 million in 2006 (137). Over the same period, the number of Medicare claims for facet joint injections increased by 76%. They also found that 63% of facet joint injection services allowed by Medicare in 2006 did not meet the Medicare program requirements, resulting in approximately \$96 million in improper payments. In addition, Medicare also allowed an additional \$33 million in improper



payments for associated facility claims. Of the error rate, 38% of facet joint injection services had a documentation error, 31% a coding error, 8% did not establish medical necessity, and 14% had overlapping errors. They also showed that approximately 50% of the procedures and most of the errors in coding were performed by non-interventional pain physicians.

### 3.1 Importance

Many of the causes of spinal pain and other chronic pain conditions are considered to be either acute recurrent problems characterized by periods of quiescence punctuated by flare-ups, or chronic diseases, like diabetes or hypertension, requiring long-term treatment with ongoing care. The importance of interventional techniques in managing chronic spinal pain has been established on the basis of advances in imaging, neuroanatomic findings, new discoveries in chemical mediation, the development of precision diagnostic and therapeutic injection techniques, and reported non-operative treatment successes. Many

guidelines, systematic reviews, Cochrane Reviews, and other articles pertaining to interventional pain management have been published (7,10-26,34-63,65-97). However, most of these guidelines are ambiguous and may not be applicable in managing chronic spinal pain utilizing contemporary interventional pain management. Further, there are quality issues with inclusion or exclusion of significant literature such as observational studies.

Thus, the quality of systematic reviews, guidelines, and policies has been questioned, and concerns have been raised regarding non-applicability across populations, bias, and alleged major short comings with potentially harmful health care implications for patients in the United States (28-33,60-63,148-154). As a result, ASIPP has developed an ongoing strict process of evidence synthesis and guideline preparation with appropriate updating since 1999 (22-26,59). The interventional techniques guidelines and opioid guidelines developed by ASIPP have been listed on the Agency for Healthcare Research and Quality (AHRQ)/National Guideline Clearinghouse (NGC) web site (154,155).

### 3.2 Technology

Diagnostic and therapeutic interventional techniques in the management of chronic spinal pain have been evaluated. These include facet joint interventions, sacroiliac joint interventions, epidural injections, lumbar epidural adhesiolysis, discography and intradiscal therapies, vertebral augmentation techniques, and implantable therapies.

## 4.0 METHODOLOGY OF GUIDELINE DEVELOPMENT

---

In recent years, there have been substantial increases in the number of treatment alternatives available to providers and patients, the proportions of patients receiving interventional pain management services, the volume of studies describing the effectiveness (or ineffectiveness) of those options, guidelines, and systematic reviews. The body of available evidence is becoming more complex, conflicting, and difficult to manage for most providers. Thus, guidelines have become a key tool for comprehensively summarizing the available literature and placing it in a format accessible to interventional pain management physicians (2,156).

### 4.1 Historical Aspects

The development of clinical practice guidelines for use by practitioners, payors, patients, and others is a key strategy in promoting the use of highly effective clinical services (2). However, the use of practice guidelines in medicine is not new (3,6,7,157-169). In fact, nearly all generic rules, treatments, indications, and criteria in medical text books can be considered as practice guidelines. The first guidelines, developed in the 1840s, shortly after the use of anesthesia was first demonstrated, were specific to the practice of anesthesia and concerned themselves with overall patient safety and effective technique for preventing anesthetic mishaps (164-167). Subsequent guidelines included the American Academy of Pediatrics' Redbook of Infectious Disease, published in the 1930s (168). Since then, numerous guidelines and practice parameters have been developed and used in the practice of medicine in the United States. The lack of interest by the physician community in general and of most interventional pain physicians in particular in the disciplined appraisal of existing medical literature led to a dependence on the teachings of senior instructors and anecdotal experiences and has resulted in variable perceptions and interpretations, and highly variable outcomes for interventions. Consequently, it has been alleged that

there has been significant unnecessary care and fraud and abuse along with over utilization of procedures in the United States resulting in escalating health care costs (107,127,133-139).

### 4.2 Essentials of Guideline Development

The Committee on Reviewing Evidence to Identify Highly Effective Clinical Services Board on Health Care Services (170) provided a roadmap for the nation in *Knowing What Works in Health Care*. This work provided 3 essential functions – priority setting, evidence assessment (systematic review), and development standards for clinical practice guidelines.

Ideally, the specific clinical recommendations that are contained within practice guidelines have been systematically developed by panels of experts who have access to the available evidence, sufficient time to absorb the information, use relevant research methods, and understand the clinical problem. These experts then make considered recommendations, which become readily available to the public in an easily understandable format. These panels are expected to be objective and to produce recommendations that are unbiased, up-to-date, and free from conflict of interest (2).

Sniderman and Furberg (1) after examining the preparation of guidelines recommended a reformed process:

- ◆ First, the requisite membership of guideline groups should be defined and include the expertise relevant to that discipline plus epidemiologists, statisticians, and experts in health care policy.
- ◆ Second, the largest part of the guideline committee membership, and in particular the leaders, should be changed from one edition to the next and each edition of the guideline should include an expiration date.
- ◆ Third, reports should not be issued unanimously unless all members fully agree to all sections. Consequently, alternate interpretations and viewpoints should be recorded and issued along with the majority opinions.
- ◆ Fourth, posting an almost final version on the Internet and inviting commentary is an attractive model. This helps ensure that there is an opportunity for exchange where legitimate differences of scientific opinion exist, before final decisions are taken.
- ◆ Fifth, before publication, guidelines should undergo independent scientific review. The journal editor should present the criticisms and suggestions that result from the reviewers to the panel

for its responses, and may require revision of the guideline document, as appropriate. The editor also should consider co-publishing alternate points of view as necessary.

- ◆ Sixth, all financial relationships with industry should be disclosed in detail, including amounts received, and should be publicly available. Receipt of substantial benefits from any company or series of companies whose products might be under consideration should disqualify that individual from participation in any guideline decision-making process. Potential and actual financial benefits gained by the authors for the next 2 years should also be limited and disclosed to the association that sponsored the guideline process.
- ◆ Seventh, associations that sponsor and promote guidelines should create joint codes to govern conflict of interest, both on the part of participants in the guideline process and the associations. Associations should not accept money from industry to sponsor, underwrite, or promote guidelines.

They concluded that the guideline development process is complex and incomplete. Further, when the evidence warrants, guidelines should respect diversity of views. In addition, guidelines must be directed only to the interests of patients and not to those who profit from them. However, they also stated that failure to reform the guideline process risks replacing one authority-based system with another, whereas the core objective should be to strengthen an evidence-based approach to improve clinical care.

#### 4.3 Guidance on Guideline Development

Field and Lohr (3) provided the medical community of the United States with a treatise of directions for the preparation of clinical practice guidelines. This document published by the National Academy of Sciences was based on a congressional mandate of November 1989, with the creation of the Agency for Healthcare Policy and Research (AHCPR-Public Law 101.239), under the broad responsibility for supporting research, data development, and other activities that will “enhance the quality, appropriateness, and effectiveness of health care services . . .” and AHCPR’s request for advice from the IOM on how it might approach practice guidelines. This comprehensive report encourages standardization and consistency in guidelines development, whether such development is undertaken independently by medical societies, other organizations, or government agencies.

Field and Lohr (3) also described that conflicts in terminology and technique characterize the field of guidelines. Such conflicts are notable for the confusion they create and for what they reflect about differences in values, experiences, and interests among different parties. As public and private guideline development activities continue to multiply, the means for coordinating these efforts continue to be limited. Coordination is necessary to resolve inconsistencies, fill in gaps, track applications and results, and assess soundness of particular guidelines. Field and Lohr (3) concluded that more and disproportionate attention is paid to developing guidelines than to implementing or evaluating them. They also provided clear definitions consistent with customary, professional, and legislative usage acceptable to important interests.

Significant diversity in the development and practice of clinical practice guidelines continues. Inconsistencies among guidelines can arise from variations and values, tolerance for risks, preferences, expertise, and conflicts of interest. Multiple professionals may simply differ in how they perceive different health outcomes and how they judge when benefits outweigh harms enough to make a service worth providing to one group and not another (3).

Eden, Wheatley, McNeil, and Sox (7) in *Knowing What Works in Health Care: A Roadmap for the Nation* from the IOM of the National Academies, updated the work by Field and Lohr (3). They concluded that even though guideline developers have adopted strategies to improve the reliability and trustworthiness of the information they provide, it is not yet possible to say that the development of clinical guidelines is based on a scientifically validated process (171). The key challenges stem from the fact that guideline development frequently forces organizations to go beyond available evidence to make practical recommendations for use in everyday practice. In view of the gaps in the evidence base that frequently exist and the variable quality of the information that is available, it has been suggested that one criterion of an effective guideline process is to have 2 separate grading systems: one for the quality of evidence and another for the recommendations themselves. Even then, there may be different interpretations about what the evidence means for clinical practice. Different interpretations can be due, for example, to conflicting viewpoints about which outcomes are the most important or which course of action is appropriate given that evidence is imperfect.



#### 4.4 Key Players of Guideline Development

Many groups produce clinical practice guidelines and recommendations. The NGC website (8) in collaboration with AHRQ (9) currently includes guidelines from 326 organizations. Medical professional societies are the most common sponsors of guidelines. In addition, patient advocacy groups, payors, government agencies, and multiple for-profit groups in the United States develop guidelines with or without conducting systematic reviews, thus, providing highly variable recommendations. The NGC included approximately 650 guidelines in 1999 and has grown to over 2,400 guidelines today. The website receives an estimated 1.3 million visits per month. For a guideline to be included on the website, guideline producers are required to demonstrate that they performed a systematic literature search and that they developed, reviewed, or revised the guideline within the last 5 years. Thus, by meeting NGC standards and being admitted to the website, guideline developers are able to improve dissemination. ASIPP guidelines have been listed on the NGC website since 2002.

Thus, the field of guidelines development is a complex and confusing arena with high expectations, competing organizations, conflicting philosophies, and ill-defined or incompatible objectives (7,22-26,59-63,77,80,97,158,172-174). The field of guidelines suffers from imperfect and incomplete scientific knowledge, as well as imperfect and uneven means of applying that knowledge. Despite the good intentions of many involved organizations and parties, guidelines development continues to lack clearly articulated goals, coherent structure, and credible mechanisms for evaluating, improving, and coordinating guidelines development to meet social needs for high quality, affordable health care.

Substantial guidance has been provided for the critical appraisal of evidence to prepare guidelines (172,175-178). Standardized approaches have also been developed to evaluate the development and validity of guidelines (148,173,174-194). However, the results of evaluations of various guidelines have been less than optimal (60-63,179).

#### 4.5 Status of Guidelines in Interventional Pain Management

Manchikanti et al (62) performed a critical appraisal of the 2007 American College of Occupational and Environmental Medicine (ACOEM) Practice Guidelines for Interventional Pain Management utilizing

the Appraisal of Guidelines Research and Evaluation (AGREE), American Medical Association (AMA), IOM, and other criteria. Critical appraisal utilizing the AGREE instrument found that both chapters scored less than 10% in 3 of the 6 domains, less than 20% in one domain, over 30% in one domain, and over 70% in one domain. Global assessment also scored below 30% with the recommendation from AGREE, "not recommended as suitable for use in practice." Further, this analysis showed that both chapters of ACOEM guidelines (10,11) met only one of the 6 AMA key attributes, only 3 of the 8 IOM criteria attributes, and only 28% of the criteria described by Shaneyfelt et al (179). The authors of the reassessment (60-62) concluded that both the low back pain and chronic pain chapters of the ACOEM guidelines may not be ideal for clinical use based on the assessment by the AGREE instrument, AMA attributes, and the criteria established by Shaneyfelt et al. Manchikanti et al (61) also reviewed the potential implications of occupational medicine practice guidelines for interventional pain management and concluded that ACOEM guidelines for interventional pain management have no applicability in modern patient care due to the lack of expertise by the developing organization (ACOEM), the lack of utilization of appropriate and current EBM principals, and the lack of significant involvement of experts in these techniques, resulting in a lack of clinical relevance. Further, they added that these guidelines may result in reduced medical quality of care; may severely hinder access to appropriate medically needed and essential medical care; and finally, they may increase costs for injured workers, third party payors, and the government by transferring the injured worker into a non-productive disability system.

Helm (63) evaluated Occupational Medicine Practice Guidelines utilizing Shaneyfelt et al's criteria (179). He concluded that ACOEM complied with only 12 out of 25 criteria. Shaneyfelt et al (179) assessed the quality of 279 guidelines produced over the period of 1985 to 1997 and assessed their quality against a set of 25 standards. The investigators found that the mean number of quality standards satisfied over that period was 11 or 43%. For example, less than 10% of the guidelines described formal methods of combining scientific evidence and expert opinion. The investigators also evaluated the guidelines in accordance with their specification of purpose (75% compliance), definition of the patient population involved (46%), pertinent health outcomes (40%), method of external review (32%),

and whether an expiration date or scheduled update was included (11%). Although they found significant improvement over time, each guideline still only met 50% of the standards, on average, in 1997.

Many of the criticisms directed at the U.S. system of guideline production in 1990 still apply today. These criticisms focused on conflicting clinical recommendations; failure to address certain topics; and incomplete public disclosure of the evidence surveyed, methods used, composition of the panel, and conflicts of interest. Aside from the role that AHRQ plays in populating the NGC website, no independent entity exists in the United States to certify guideline quality or to develop national standards regulating the content or methods of guideline developers.

Multiple criteria have been described. The AGREE instrument (186), Shaneyfelt et al (179), National Health and Medical Research Council (NHMRC) (182), World Health Organization (WHO) (195), AMA (196), CMA (197), and AHCP (198) have all carefully formulated the methodology for developing scientifically sound guidelines and rating of the strength of evidence.

#### **4.6. Controversies in Guideline Development**

Clinical guidelines in the United States are developed by government agencies, medical professional societies, patient advocacy groups, trade associations, and multiple other for-profit organizations. Many consider the U.S. Preventive Services Task Force (USPSTF) to be a model for clinical recommendations development (7). USPSTF conducts impartial assessments with scientific evidence to reach conclusions about the effectiveness of a broad range of clinical preventive services, including screening, counseling, and preventive medications.

USPSTF recommendations are based on systematic reviews of the evidence on specific topics on clinical prevention; these reviews are performed by Evidence-Based Practice Centers (EPCs) (199). It has been generally stated that the guidelines developed by government agencies reflect the fact that the production of high quality guidelines requires substantial and sufficient resources and that government agencies have more resources available to do the work (187,200). However, these reviews may be performed by related organizations, with the introduction of unintended bias and conflicts. Even bearing this in mind, the best accepted guidelines tend to be ones from the USPSTF, the National Institute for Clinical Excellence (NICE) in the United Kingdom, and WHO.

Guidelines often elicit controversy for numerous reasons, including the type of recommendations and the restrictions on practice patterns. In fact, Congress eliminated the AHCP in 1995 soon after the development of acute low back pain guidelines (201). The AHCP, over the years, issued 19 guidelines at a cost of \$750 million (over \$40 million per guideline). Those guidelines were not demonstrated to have saved health care dollars and were not widely utilized, thus questioning the cost effectiveness of governmentally developed guidelines (77,201).

Smaller professional organizations are considered to lack the internal resources, including staff capacity and expertise, required to produce guidelines (2). It has been stated that this is true when the organization produces both systematic reviews and the guideline recommendations, 2 tasks supposedly requiring different skill sets. Further, it is stated that even larger professional organizations can face resource constraints in this area. However, the lack of utilization of governmental agency produced guidelines in the United States due to the private health care system, their expense, and the bureaucracy of larger organizations (similar to the government) raises numerous questions on these assumptions. Further, in subjects without extensive literature, the role of methodologists may be an exercise in futility.

Conflicts of interest in guideline development and inappropriate methodologies have been questioned. The impropriety is based on pharmaceutical and medical device company sponsorship, when members of the guidelines committees have a substantial financial association with an industry, when there is a relationship between the developing organization and industry, or, finally, when there is no relevant clinical relationship or expertise on the part of the developers of the guidelines (1,3,7,8,10-13,30,61,62,106,144,148,179,182,187,196,202-207). The financial ties between guideline panels and industry appear to be extensive. In fact, a survey of 685 disclosure statements by authors of guidelines concerning medications found that 35% declared a potential financial conflict of interest (187).

Sniderman and Furberg (1) extensively discussed the conflicts, controversies, and limitations of the guideline process. They considered the anchoring authority of the guideline process is the belief that guidelines are evidence-based, not opinion based, and therefore their conclusions flow directly from the conclusions of the studies. Accordingly, the outcome is perceived to be impersonal and inevitable. Limitations of the guide-

lines process include governance and composition of the guideline committee, unanimity in guidelines, lack of independent review, and conflict of interest.

Even though fundamental principles and operating procedures have been suggested for the guideline process, most of the guideline committees have considerable latitude to establish their own working rules. In fact, this appears to be the major flaw in the development of guidelines, in EBM synthesis, and guideline development in general and in interventional pain management in particular. In preparation of the medical guidelines, there do not appear to be explicit rules as to the range of the expertise that must be included within the committee with minimal representation of epidemiologists, economists, and sometimes the experts themselves. Even though in general, guidelines strive to be evidence-based, or at least the focus is evidence-based, they cannot be derived strictly, solely, and incontestably from the evidence. On many issues, if not most, the evidence, no matter how extensive, remains incomplete (1).

In general, in medicine, unanimity in guidelines is an exception rather than the rule. Generally it is a tactic, not a necessary result (1). Consequently, unanimity is strikingly absent when different guidelines are compared. The next issue is with regards to lack of independent review, followed by conflict of interests. By favoring one test over another, or one therapy over another, guidelines often create commercial winners and losers, who cannot be disinterested in the result and who therefore must be separated from the process (1). Unfortunately, the groups that finance medical care do not automatically accept the recommendations of guidelines. Similarly, those who write the guidelines and those who issue them also have significant conflict of interest (208,209).

#### **4.7 ASIPP Guideline Development Process**

ASIPP launched the development of practice guidelines for interventional techniques in the management of chronic pain in 1999 and published the first guideline in 2000 (26). These guidelines were started to create a document to help practitioners by synthesizing the available evidence. The authors stated that these clinical practice guidelines for interventional techniques in the management of chronic pain were professionally developed utilizing a combination of evidence and consensus.

The synthesis of evidence, committee composition, and the development process have been revised,

refined, and expanded with evaluation at least once every 2 years (22-48,210-216).

ASIPP guidelines meet most criteria described by Shaneyfelt et al (179), AGREE (186), IOM (3), as well as the majority of recommendations by Sniderman and Furberg (1).

### **5.0 DEVELOPMENT OF GUIDELINES**

The era of the physician as the sole health care decision-maker is long passed. In the modern world, health care decisions are made by multiple entities and personnel, individually or in collaboration, in multiple contexts for multiple purposes. The decision-maker is likely to be the payor or the intermediary employed by the payor, or the consumer himself. Consequently, every decision-maker needs credible, unbiased, and understandable evidence on the effectiveness of health interventions and services, in addition to the physician.

The guideline development process includes evidence assessment through systematic reviews and developing clinical practice guidelines with recommendations.

#### **5.1 Evidence-Based Medicine**

Historically, EBM is traceable to the 1700s, even though it was not explicitly defined and advocated until the late 1970s and early 1980s (4). Initially, EBM was called "critical appraisal" to describe the application of basic rules of evidence. Thus, EBM is about solving clinical problems (98) and acknowledges that intuition, unsystematic clinical experience, and pathophysiologic rationale are insufficient grounds for clinical decision-making, and it stresses the examination of evidence from clinical research (100), in contrast to the traditional paradigm of medical practice. EBM suggests that a formal set of rules must complement medical training and common sense for clinicians to interpret the results of clinical research effectively. Thus, EBM requires the prudent, specific contextual application of knowledge gained by integration of individual clinical expertise and experience, in concert with the best available external evidence gained from systematic research (29-33,217).

EBM explicitly mandates the necessity for an active search for all information that is valid and relevant, in an ongoing assessment to ascertain both the accuracy of information and the applicability of evidence to the decision in question. Evidence-based practice emphasizes an integration of the best research evidence with

the patient's circumstances and values. Thus, an ethical and practical approach to EBM involves 2 fundamental principles. First, delineating that scientific evidence alone is never sufficient to make a clinical decision. Decision-makers must always consider the patient's values when evaluating the benefits, risks, and burdens associated with any and all treatment strategies (98). The second principle is that EBM describes a hierarchy of informational value to guide clinical decision-making (98,100). This hierarchy is never absolute. EBM must reflect how different types and levels of evidence can be relative to and inform the calculus of circumstances, agents, and the consequences of decisions and actions (217).

Evidence-based practice is defined based on 4 basic and important contingencies (218):

- ◆ Recognition of the patient's problem and construction of a structured clinical question.
- ◆ Thorough search of the medical literature to retrieve the best available evidence to answer the question.
- ◆ Critical appraisal of all available evidence.
- ◆ Integration of the evidence with all aspects and contexts of the clinical circumstances to facilitate the decisional process that determines the best clinical care of each patient.

## 5.2 Hierarchy of Strength of Evidence

A hierarchy of strength of evidence for treatment decisions provided by Guyatt and Drummond (99) is as follows:

- ◆ N of 1 randomized controlled trial
- ◆ Systematic reviews of randomized trials
- ◆ Single randomized trial
- ◆ Systematic review of observational studies addressing patient-important outcomes
- ◆ Single observational study addressing patient-important outcomes
- ◆ Physiologic studies (studies of blood pressure, cardiac output, exercise capacity, bone density, and so forth)
- ◆ Unsystematic clinical observations

The N of 1 RCT, which is at the top of the hierarchy of strength of evidence for treatment decisions, is one in which patients undertake pairs of treatment periods receiving a target treatment during one period of each pair and a placebo or alternative during the other. Patients and clinicians are blind to allocation, the order of the target and control is randomized, and patients make quantitative ratings of their symp-

oms during each period. The N of 1 RCT continues until both the patient and clinician conclude that the patient is, or is not, obtaining benefit from the target intervention. N of 1 RCTs are applicable to reversible therapies, such as mood altering medications. N of 1 RCTs are not applicable to most surgeries for several reasons. Once the procedure is done it may not be undone. The ethics of providing sham surgeries for blinding is highly questionable. The expense of such studies is often prohibitive.

The philosophy of a hierarchy of evidence in guiding therapy, though not absolute, has created emphasis on the importance of randomized trials. In modern medicine, most researchers synthesizing the evidence may or may not follow the rules of EBM, which requires that a formal set of rules must complement medical training and common sense for clinicians to interpret the results of clinical research. However, guideline developers and systematic reviewers seem to ignore that very different hierarchies are necessary for each use of therapy, diagnosis, and prognosis (219).

Most questions in medical research are investigated in observational studies which are infrequently used in evidence synthesis by some organizations including the Cochrane Review group (220-225). Observational studies are more likely to provide an indication of daily medical practices (225). Consequently, the proponents of observational studies proclaim that observational studies are just as effective as RCTs. From a methodologic perspective, the 2 types of studies are considered complementary rather than opposing (224). In the setting of modern clinical research, observational studies and RCTs can be viewed as expressions of the steps of observation and experimentation that form the basis of scientific methodology. It has been described that the observational step is used to uncover patterns and formulate hypotheses regarding cause-and-effect relationships. In the experimentation step the observational hypotheses are confirmed or refuted in an experiment in which the independent variables are controlled by the experimenter (224,226,227). Consequently, the argument about one or other evidence is misplaced since both observation and experimentation steps are required for scientific advancement (224).

## 5.3 Level of Evidence

The translation of systematic reviews into practice recommendations is not straightforward. The same information can be interpreted in different

ways by different panelists, resulting in the provision of different guidance (228). Often, even when there is substantial consensus about what the scientific evidence says, there are disagreements about what the evidence means for clinical practice. Conclusions about clinical effectiveness can vary widely as a result of conflicting viewpoints, such as which outcomes are the most important and which course of action is appropriate given that the evidence is imperfect. Thus, systematic reviews assess the quality of the individual studies and provide the quality and level of evidence. In developing guidelines both are important, that is the quality of evidence and the strength of recommendation which takes into account the balance of the benefits and harms that are associated with the intervention.

The evidence base that supports clinical practice guidelines is often quite limited and guideline developers must often wrestle with what to do when “the irresistible force of the need to offer clinical advice meets with the immovable object of flawed evidence”

(2,229,230). The authors of guidelines must consider the best way to address the trade-off between rigor and pragmatism, and between adherence to evidence and broader clinical utility (2,231,232). The authors may nonetheless state their evaluation and recommendations based upon the current best available evidence.

### 5.3.1 Determination of Level of Evidence

Level of evidence is derived from quality assessment and the results of individual studies. While there is no universally accepted approach to presenting levels of evidence, a rigorous approach in widespread use was developed by the USPSTF (233). Table 1 illustrates the quality of evidence developed by the USPSTF. However, most guidelines and systematic reviews utilized the hierarchy of evidence reported by the AHRQ (formerly the AHCPR hierarchy of evidence, funding for which was eliminated by Congress) and this guideline also carries the disclaimer “not for patient care,” as shown in Table 2 (77).

Table 1. *Quality of evidence developed by USPSTF.*

<b>I:</b>	<b>Evidence obtained from at least one properly randomized controlled trial</b>
<b>II-1:</b>	<b>Evidence obtained from well-designed controlled trials without randomization</b>
<b>II-2:</b>	<b>Evidence obtained from well-designed cohort or case-control analytic studies, preferably from more than one center or research group</b>
<b>II-3:</b>	<b>Evidence obtained from multiple time series with or without the intervention. Dramatic results in uncontrolled experiments (such as the results of the introduction of penicillin treatment in the 1940s) could also be regarded as this type of evidence</b>
<b>III:</b>	<b>Opinions of respected authorities, based on clinical experience descriptive studies and case reports or reports of expert committees</b>

Adapted from the U.S. Preventive Services Task Force (USPSTF) (233).

Table 2. *Panel ratings of available evidence supporting guideline statements.*

A = Strong research-based evidence (multiple relevant and high-quality scientific studies).
B = Moderate research-based evidence (one relevant high-quality scientific study or multiple adequate scientific studies*).
C = Limited research-based evidence (at least one adequate scientific study* in patients with low back pain).
D = Panel interpretation of information that did not meet inclusion criteria as research-based evidence.

\* Met minimal formal criteria for scientific methodology and relevance to population and specific method addressed in guideline statement.

Source: Bigos SJ et al. Acute Low Back Problems in Adults. Clinical Practice Guideline No. 14, AHCPR Publication No. 95-0642. Rockville, Maryland, U.S.A., Agency for Health Care Policy and Research, Public Health Service, U.S., Department of Health and Human Services, December, 1994 (77).

In an AHRQ commissioned document with overreaching goals of describing systems to rate the strength of scientific evidence, West et al (172) in their evidence report/technology assessment of Systems to Rate the Strength of Scientific Evidence also included evaluation of the quality of individual articles that make up a body of evidence on a specific question in health care and provided guidance as to best practices. For determination of the level of evidence, methodologic quality is crucial to which all aspects of study design and conduct can be shown to protect against systematic bias, non-systematic bias, and inferential error.

Multiple systems also have been developed by Cochrane Review Group and others. These are well described in numerous other documents (22,30-33,60-63,68,76,172). West et al (172) described quality as the aggregate of quality ratings for individual studies, predicated on the extent to which bias was minimized. Quantity is the magnitude of effect, numbers of studies, and sample size of power. Consistency is that for any given topic the extent to which similar findings are reported using similar and different study designs. They reviewed 40 systems that grade the strength of a body of evidence: 34 from sources other than AHRQ EPCs and 6 from EPCs based on quality, quantity, and consistency.

Systems for grading the strength of a body of evidence are much less uniform than those for rating study quality. West et al (172), despite being able to identify various rating and grading systems that can more or less be taken off the shelf for use, found many areas in which information or empirical documentation was lacking, and provided recommendations for future research. They also cautioned that until these research gaps are bridged, those wishing to produce authoritative systematic reviews or technology assessments will be somewhat hindered in this phase of their work. They specifically highlighted the need for work on: 1) identifying and resolving quality rating issues pertaining to observational studies; 2) evaluating inter-rater reliability of both quality rating and strength-of-evidence grading systems; 3) comparing the quality ratings from different systems applied to articles on a single clinical or technology topic; 4) similarly, comparing strength-of-evidence grades from different systems applied to a single body of evidence on a given topic; 5) determining what factors truly make a difference in final quality scores for individual articles (and by extension a difference in how quality is

judged for bodies of evidence as a whole); 6) testing shorter forms in terms of reliability, reproducibility, and validity; 7) testing applications of these approaches for "less traditional" bodies of evidence such as systematic reviews of disease risk factors, screening tests, and counseling interventions; 8) assessing whether the study quality grids that were developed by AHRQ are useful for discriminating among studies of varying quality and, if so, refining and testing the systems further using typical instrument development techniques; and finally, 9) comparing and contrasting approaches to rating quality and grading evidence strength in the United States and abroad.

Methodologic quality assessment of systematic reviews is crucial for guideline preparation and grading recommendations. West et al (172) described a set of high-performing scales or checklists pertaining to systematic reviews, with 7 key domains: study question, search strategy, inclusion and exclusion criteria, data abstraction, study quality and validity, data synthesis and analysis, and funding or sponsorship (Table 3).

#### 5.4 Systematic Reviews

Systematic reviews are central to scientific inquiry into what is known and not known about what works in health care (234). A systematic review is a scientific investigation that focuses on a specific question and uses explicit, preplanned scientific methods to identify, select, assess, and summarize similar but separate studies. It may or may not include meta-analysis which is a quantitative synthesis.

Fundamentals of a systematic review include 5 basic steps:

- Step 1 – formulate the research question;
- Step 2 – construct an analytic (or logic) framework;
- Step 3 – conduct a comprehensive search for evidence;
- Step 4 – critically appraise the evidence; and
- Step 5 – synthesize the body of evidence.

With Step 1, a well formulated question guides the analytic (or logic) framework for the review, the overall research protocol, and the critical appraisal of the relevant evidence.

Step 2 involves construction of an analytic framework. Once the research question is established, it should be articulated in an analytic framework that clearly lays out the chain of logic underlying the case for the health intervention of interest. The complexity of the analysis will vary depending on the number of linkages between the intervention and the outcomes of interest. The choice of study designs to be included

Table 3. *Domains in the Agency for Healthcare Research and Quality (AHRQ) criteria for evaluating systematic reviews.*

DOMAIN	ELEMENTS*
Study question	<ul style="list-style-type: none"> <li>• <b>Question clearly specified and appropriate</b></li> </ul>
Search strategy	<ul style="list-style-type: none"> <li>• <i>Sufficiently comprehensive and rigorous with attention to possible publication biases</i></li> <li>• <i>Search restrictions justified (e.g., language or country of origin)</i></li> <li>• Documentation of search terms and databases used</li> <li>• Sufficiently detailed to reproduce study</li> </ul>
Inclusion and exclusion criteria	<ul style="list-style-type: none"> <li>• <b>Selection methods specified and appropriate, with <i>a priori</i> criteria specified if possible</b></li> </ul>
Interventions	<ul style="list-style-type: none"> <li>• <b>Intervention(s) clearly detailed for all study groups</b></li> </ul>
Outcomes	<ul style="list-style-type: none"> <li>• <b>All potentially important harms and benefits considered</b></li> </ul>
Data extraction †	<ul style="list-style-type: none"> <li>• Rigor and consistency of process</li> <li>• Number and types of reviewers</li> <li>• Blinding of reviewers</li> <li>• Measure of agreement or reproducibility</li> <li>• Extraction of clearly defined interventions/exposures and outcomes for all relevant subjects and subgroups</li> </ul>
Study quality and validity	<ul style="list-style-type: none"> <li>• <i>Assessment method specified and appropriate</i></li> <li>• Method of incorporation specified and appropriate</li> </ul>
Data synthesis and analysis	<ul style="list-style-type: none"> <li>• <i>Appropriate use of qualitative and/or quantitative synthesis, with consideration of the robustness of results and heterogeneity issues</i></li> <li>• Presentation of key primary study elements sufficient for critical appraisal and replication</li> </ul>
Results	<ul style="list-style-type: none"> <li>• <b>Narrative summary and/or quantitative summary statistic and measure of precision, as appropriate</b></li> </ul>
Discussion	<ul style="list-style-type: none"> <li>• <b>Conclusions supported by results with possible biases and limitations taken into consideration</b></li> </ul>
Funding or sponsorship	<ul style="list-style-type: none"> <li>• <i>Type and sources of support for study</i></li> </ul>

\* Elements appearing in italics are those with an empirical basis. Elements appearing in bold are those considered essential to give a system a Yes rating for the domain.

† Domain for which a Yes rating required that a majority of elements be considered.

Source: West S et al. Systems to Rate the Strength of Scientific Evidence, Evidence Report, Technology Assessment No. 47. AHRQ Publication No. 02-E016. Rockville, MD: Agency for Healthcare Research and Quality, 2002. [www.thecre.com/pdf/ahrq-system-strength.pdf](http://www.thecre.com/pdf/ahrq-system-strength.pdf) (172).

in a systematic review should be based on the type of the research question being asked and should have the goal of minimizing bias. RCTs can answer questions about the efficacy of screening, preventive, and therapeutic interventions. Even though RCTs can best answer questions about the potential harms from interventions, observational study designs, such as cohort studies, case series, or case control studies, may be all that are available or possible for the evaluation of rare or long-term outcomes (234). In fact, because harms from interventions are often rare or occur far in the future, a systematic review of observational research may be the best approach to identifying reliable evidence on potential rare harms or benefits. However, observational studies are generally the most appropriate for answering questions related to prognosis, diagnostic accuracy, incidence, prevalence, and

etiology (234-236). Cohort studies and case series are useful for examining long-term outcomes because RCTs may not monitor patients beyond the primary outcome of interest or for rare outcomes because they generally have small numbers of participants.

Thus, systematic reviews performed for interventional studies utilized RCTs and observational studies. However, for interventional procedures as well as surgical procedures, population-based public health measures, quality improvement strategies, and many other health care interventions, relevant, randomized evidence is frequently unavailable (237,238). Therefore, it is critical to utilize all types of evidence appropriately for therapeutic and diagnostic accuracy studies. Indeed, the evidence base for the effectiveness of most health services is sparse. Well-designed, well-conducted studies of the effectiveness of most

health care services are the exception. The available research evidence falls far short of answering many questions that are important to patients and providers (239). In a review of the quality of evidence for 69 medical devices, surgical procedures, and other medical therapies that were subject to national Medicare coverage determinations from 1998 to 2003 (240), the researchers found good evidence on health outcomes for only 11 of the 69 technologies (16%). Further, for more than 29 technologies, there was either no evidence at all (6 technologies) or poor-quality evidence (23 technologies). Poor quality was due to a limited number of studies, the weak power of the studies, flaws in the design or the conduct of the studies, or missing information on important health outcomes. The evidence was considered "fair" only for 42% of the technologies.

Step 3 involves conducting a comprehensive search for evidence. It is arguably the most important step in conducting a high-quality systematic review. It is crucial to identify all the relevant studies meeting the eligibility criteria for the systematic review. A comprehensive search is necessary because there is no way

of knowing whether the missing studies are missing at random or missing for a reason critical to understanding current knowledge. However, most systematic reviews often miss significant portions of the literature, take one to 2 years to prepare, and 6 months to 2 years for publication. A comprehensive search strategy is essential including hand searching.

Step 4 involves critically appraising the evidence, assessing the study quality, strength of findings, consistency, external validity, and estimate of the effect. Both experimental and observational studies must be judged for their external validity and for their applicability to the population of interest.

Multiple systems are available for the quality assessment of randomized trials, observational studies, and diagnostic accuracy studies (68,76,172,120-122,124,125,238). Tables 4 - 6 illustrate commonly utilized methodologic quality assessment instruments.

### 5.5 Grading Recommendations

Strength of evidence and grading recommendations are not the same, but are crucial in the clini-

Table 4. *Modified AHRQ methodologic assessment criteria for diagnostic interventions.*

CRITERION	Weighted Score (points)
<b>1. Study Population</b>	<b>15</b>
• Subjects similar to populations in which the test would be used and with a similar spectrum of disease	
<b>2. Adequate Description of Test</b>	<b>10</b>
• Details of test and its administration sufficient to allow for replication of study	
<b>3. Appropriate Reference Standard</b>	<b>30</b>
• Appropriate reference standard (gold standard) used for comparison	15
• Reference standard reproducible	15
<b>4. Blinded Comparison of Test</b>	<b>30</b>
• Evaluation of test without knowledge of disease status, if possible	15
• Independent, blind interpretation of test and reference	15
<b>5. Avoidance of Verification Bias</b>	<b>15</b>
• Decision to perform reference standard not dependent on results of test under study	
TOTAL SCORE	100

Adapted and modified from West S et al. *Systems to Rate the Strength of Scientific Evidence, Evidence Report*, Technology Assessment No. 47. AHRQ Publication No. 02-E016 (172).



Table 5. *Modified AHRQ quality assessment criteria for observational studies.*

CRITERION	Weighted Score (points)
<b>1. Study Question</b>	<b>2</b>
• Clearly focused and appropriate question	
<b>2. Study Population</b>	<b>8</b>
• Description of study population	5
• Sample size justification	3
<b>3. Comparability of Subjects for All Observational Studies</b>	<b>22</b>
• Specific inclusion/exclusion criteria for all groups	5
• Criteria applied equally to all groups	3
• Comparability of groups at baseline with regard to disease status and prognostic factors	3
• Study groups comparable to non-participants with regard to confounding factors	3
• Use of concurrent controls	5
• Comparability of follow-up among groups at each assessment	3
<b>4. Exposure or Intervention</b>	<b>11</b>
• Clear definition of exposure	5
• Measurement method standard, valid and reliable	3
• Exposure measured equally in all study groups	3
<b>5. Outcome measures</b>	<b>20</b>
• Primary/secondary outcomes clearly defined	5
• Outcomes assessed blind to exposure or intervention	5
• Method of outcome assessment standard, valid and reliable	5
• Length of follow-up adequate for question	5
<b>6. Statistical Analysis</b>	<b>19</b>
• Statistical tests appropriate	5
• Multiple comparisons taken into consideration	3
• Modeling and multivariate techniques appropriate	2
• Power calculation provided	2
• Assessment of confounding	5
• Dose-response assessment if appropriate	2
<b>7. Results</b>	<b>8</b>
• Measure of effect for outcomes and appropriate measure of precision	5
• Adequacy of follow-up for each study group	3
<b>8. Discussion</b>	<b>5</b>
• Conclusions supported by results with possible biases and limitations taken into consideration	
<b>9. Funding or Sponsorship</b>	<b>5</b>
• Type and sources of support for study	
<b>TOTAL SCORE</b>	<b>100</b>

Adapted and modified from West S et al. *Systems to Rate the Strength of Scientific Evidence*, Evidence Report, Technology Assessment No. 47. AHRQ Publication No. 02-E016 (172).

Table 6. *Modified weighted Cochrane methodologic quality assessment criteria.*

Criterion		Weighted Score (points)
1. Internal validity		63
A	Selection and restriction	4
B	Treatment allocation (randomization process and concealment should be provided in detail)	15
C	Prognostic comparability. The distribution of baseline characteristics is similar and clearly presented for intervention groups.	10
D	Blinding of patients	4
E	Blinding of physician	4
F	Blinding of observer	4
G	Dropouts	12
H	Loss to follow-up assessment	10
2. Relevance		20
I	Extra treatments or co-interventions similar.	2
J	Intervention Detailed description of interventions	5
K	Outcome measures	5
L	Timing of outcome measurements	6
M	Side effects	2
3. Statistical approaches		17
N	Analysis and presentation of data	5
O	Study size	12
TOTAL SCORE		100

Adapted from Koes BW et al. Efficacy of epidural steroid injections for low-back pain and sciatica: A systematic review of randomized clinical trials. *Pain* 1995; 63:279-288 (76).

cal application of guidelines (196,241-256). Over the years, many guideline programs have developed a structured methodology following "evidence-based principles" that are complemented by consensus processes. However, others are still in search of guidance (3,172,256-261). The process of appraising evidence and formulating a recommendation has been frequently described (Fig. 2) (256). However, the strength of recommendations and grading has been misinterpreted. Guidelines making recommendations always make judgments about the quality of evidence and the balance of benefits and downsides, which include harms, burdens, and costs. Frequently, these judgments are made implicitly rather than explicitly. Judgments about the quality of evidence are confused with judgments about the balance of benefits and risks. Many systems that are used to grade the quality of evidence and strength of recommendations also confuse these judgments by equating the strength of recommendation with the quality of evidence. For example it would be wrong to grade recommendations

from which there is high quality evidence as strong without explicitly considering the balance of benefits and risks. Knowing the quality of evidence is essential, but not sufficient, for making judgments about the strength of recommendation. For instance, high quality evidence from well executed RCTs showed that oral anticoagulation administered for more than one year reduces the risk for recurrent thromboembolic events in patients after a first episode of spontaneous deep venous thrombosis (249). However, because oral anticoagulation is associated with harms (bleeding risk), burden (taking medication and monitoring anticoagulation levels), and cost (anticoagulation clinics or monitoring devices), the recommendation to anticoagulate all patients is weakened; the benefits and downsides are finely balanced and individual patients will make different choices. Consequently, both judgments about the quality of evidence and about the strength of recommendation are complex and require consideration of a number of factors. For over 25 years, a growing number of organizations have

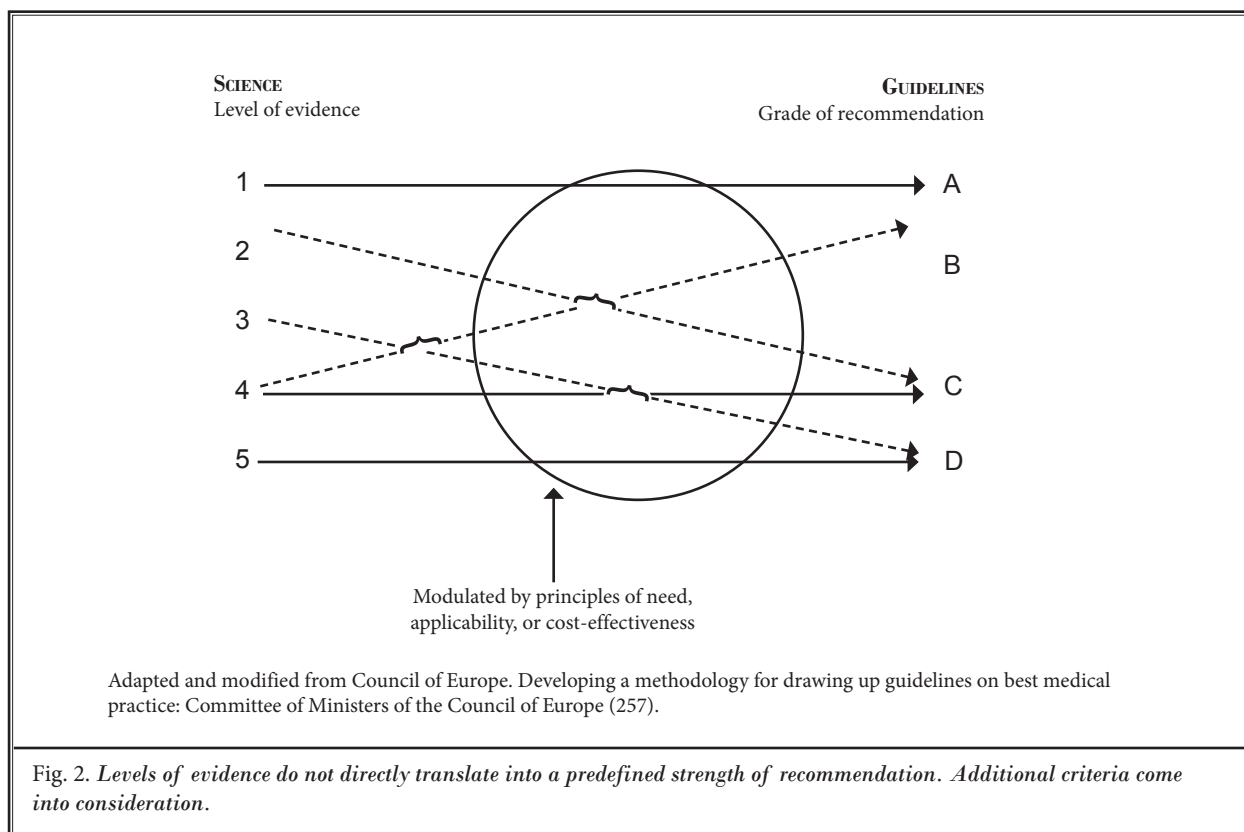


Fig. 2. Levels of evidence do not directly translate into a predefined strength of recommendation. Additional criteria come into consideration.

employed various systems to grade the quality of evidence (sometimes called levels of evidence) and the strength of recommendations (249). Unfortunately, different organizations use various grading systems, which may lead to confusion, misinterpretation, and misrepresentation, leading to deleterious effects on patient care and access (60-62).

Ebell et al (245) provided Strength of Recommendation Taxonomy (SORT): A Patient-Centered Approach to Grading Evidence in the Medical Literature in 2004. This communication was as a result of an ongoing effort in the medical publishing field to improve the quality of review articles through the use of more explicit grading of the strength of evidence on which recommendations are based (262-270). They described the problem that a Level B recommendation in one journal may not mean the same thing as Level B recommendation in another. Further, even within journals, different evidence-grading scales sometimes are used in different articles within the same issue of a journal. Since readers do not have the time, energy, or interest in multiple grading scales and more complex scales are difficult to integrate into daily practices, they devel-

oped a new taxonomy known as SORT. This taxonomy while it was meant for U.S. family medicine and primary care journals may be utilized by other journals and specialties also. The attributes of this taxonomy include uniformity; it allows authors to evaluate the strength of recommendations of a body of evidence; it rates the level of evidence for an individual study; it is comprehensive and allows authors to evaluate studies of screening, diagnosis, therapy, prevention, and prognosis; it is easy to use and not too time-consuming for authors, reviewers, and editors who may be content experts but not expert in critical appraisal or clinical epidemiology; and it is straightforward enough that physicians can readily integrate the recommendations into daily practice. Table 7 illustrates the SORT Taxonomy with ratings of A, B, or C for the strength of recommendation for a body of evidence. Table 8 explains whether a body of evidence represents good or limited-quality evidence and whether evidence is consistent or inconsistent. The quality of individual studies is rated as 1, 2, or 3. Numbers are used to distinguish ratings of individual studies from the letters A, B, and C used to evaluate the strength of

Table 7. Illustration of the Strength of Recommendation Taxonomy (SORT).

Strength of recommendation	Definition
A	Recommendation based on consistent and good-quality patient-oriented evidence.*
B	Recommendation based on inconsistent or limited-quality patient-oriented evidence.*
C	Recommendation based on consensus, usual practice, opinion, disease-oriented evidence,* or case series for studies of diagnosis, treatment, prevention, or screening.

\*—Patient-oriented evidence measures outcomes that matter to patients: morbidity, mortality, symptom improvement, cost reduction, and quality of life. Disease-oriented evidence measures intermediate, physiologic, or surrogate end points that may or may not reflect improvements in patient outcomes (e.g., blood pressure, blood chemistry, physiologic function, pathologic findings).

Source: Ebell MH et al. Strength of recommendation taxonomy (SORT): A patient-centered approach to grading evidence in the medical literature. *Am Fam Physician* 2004; 69:548-556 (245).

Table 8. Illustration of level of evidence based on the Strength of Recommendation Taxonomy (SORT).

Study Quality	Diagnosis	Treatment/prevention/screening	Prognosis
Level 1 – good-quality patient-oriented evidence	Validated clinical decision rule SR/meta-analysis of high-quality studies High-quality diagnostic cohort study†	SR/meta-analysis of RCTs with consistent findings High-quality individual RCT‡ All-or-none study§	SR/meta-analysis of good-quality cohort studies Prospective cohort study with good follow-up
Level 2 – limited-quality patient-oriented evidence	Unvalidated clinical decision rule SR/meta-analysis of lower-quality studies or studies with inconsistent findings Lower-quality diagnostic cohort study or diagnostic case-control study§	SR/meta-analysis of lower-quality clinical trials or of studies with inconsistent findings Lower-quality clinical trial‡ Cohort study Case-control study	SR/meta-analysis of lower-quality cohort studies or with inconsistent results Retrospective cohort study or prospective cohort study with poor follow-up Case-control study Case series
Level 3 – other evidence	Consensus guidelines, extrapolations from bench research, usual practice, opinion, disease-oriented evidence (intermediate or physiologic outcomes only), or case series for studies of diagnosis, treatment, prevention, or screening		

†—High-quality diagnostic cohort study: cohort design, adequate size, adequate spectrum of patients, blinding, and a consistent, well defined reference standard.

‡—High-quality RCT: allocation concealed, blinding if possible, intention-to-treat analysis, adequate statistical power, adequate follow-up (greater than 80 percent).

§—In an all-or-none study, the treatment causes a dramatic change in outcomes, such as antibiotics for meningitis or surgery for appendicitis, which precludes study in a controlled trial.

Source: Ebell MH et al. Strength of recommendation taxonomy (SORT): A patient-centered approach to grading evidence in the medical literature. *Am Fam Physician* 2004; 69:548-556 (245).

a recommendation based on a body of evidence. The authors of this communication described the SORT system as straightforward and comprehensive, yet easily applied by authors and physicians, while explicitly addressing the issues of patient-oriented versus disease-oriented evidence. SORT is distinguished from other evidence grading scales based on disease-oriented evidence. However, these strengths also create some

limitations as some clinicians may be concerned that this taxonomy is not as detailed in its assessment of study designs as others, such as that of the Centre for Evidence-Based Medicine (CEBM) (259). However, the authors note that the primary difference between the 2 taxonomies is that the CEBM version distinguishes between good and poor observational studies, whereas the SORT version does not.

Atkins et al (247) evaluated the quality of evidence and strength of recommendations in a pilot study group to examine the Grading of Recommendations, Assessment, Development and Evaluation (GRADE) system. The results showed that there was a varied amount of agreement on the quality of evidence for the outcomes relating to the use of the 12 questions. There was fair agreement about the relative importance of each outcome, while there was poor agreement about the balance of benefits and harms and recommendations. Most of the disagreements were easily resolved through discussion. In general, the authors found the GRADE approach to be clear, understandable, and sensible. They concluded that judgments about evidence and recommendations are complex with some subjectivity, especially regarding recommendations. Guyatt et al (255) published a description of grading the strength of recommendations and quality of evidence in clinical guidelines (Table 9). Subsequently, multiple manuscripts were

published in a series of articles discussing the GRADE system (251-253).

## 5.6 Process of Development of Guidelines

The literature available in assisting guideline preparation is extensive (3,173,179,182,184-186,190,196,241-244). Table 10 illustrates the essential components required for guideline development which has been derived from various commonly used evaluation instruments (AGREE, AMA, IOM, and Shaneyfelt et al's criteria) (3,179,190,196).

### 5.6.1 Sequential Process

The GRADE Working Group described a sequential process for developing guidelines as shown in Table 11. The GRADE system enables more consistent judgments, and communication of search judgements can support better-informed choices in health care. Table 12 shows a comparison of GRADE and other systems. As illustrated in Table 12, multiple systems were

Table 9. Grading recommendations.

Grade of Recommendation/ Description	Benefit vs Risk and Burdens	Methodological Quality of Supporting Evidence	Implications
1A/strong recommendation, high-quality evidence	Benefits clearly outweigh risk and burdens, or vice versa	RCTs without important limitations or overwhelming evidence from observational studies	Strong recommendation, can apply to most patients in most circumstances without reservation
1B/strong recommendation, moderate quality evidence	Benefits clearly outweigh risk and burdens, or vice versa	RCTs with important limitations (inconsistent results, methodological flaws, indirect, or imprecise) or exceptionally strong evidence from observational studies	Strong recommendation, can apply to most patients in most circumstances without reservation
1C/strong recommendation, low-quality or very low-quality evidence	Benefits clearly outweigh risk and burdens, or vice versa	Observational studies or case series	Strong recommendation but may change when higher quality evidence becomes available
2A/weak recommendation, high-quality evidence	Benefits closely balanced with risks and burden	RCTs without important limitations or overwhelming evidence from observational studies	Weak recommendation, best action may differ depending on circumstances or patients' or societal values
2B/weak recommendation, moderate-quality evidence	Benefits closely balanced with risks and burden	RCTs with important limitations (inconsistent results, methodological flaws, indirect, or imprecise) or exceptionally strong evidence from observational studies	Weak recommendation, best action may differ depending on circumstances or patients' or societal values
2C/weak recommendation, low-quality or very low-quality evidence	Uncertainty in the estimates of benefits, risks, and burden; benefits, risk, and burden may be closely balanced	Observational studies or case series	Very weak recommendations; other alternatives may be equally reasonable

Adapted from Guyatt G et al. 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 (255).

Table 10. *Illustration of the essential components required for guidelines derived from multiple evaluation instruments.*

AGREE (190)	AMA (196)	IOM (3)	SHANEYFELT ET AL (179)
Scope and Purpose	Organization	Validity	Standards of Guidelines Development and Format
1. Objective(s) of the guideline described. 2. Clinical question(s). 3. Population covered.	Practice guidelines should be developed by or in conjunction with physician organizations.	Practice guidelines are valid if, when followed, they lead to the health and cost outcomes projected for them, other things being equal.	1. Purpose. 2. Rationale and importance. 3. The participants and expertise. 4. Targeted health problem or technology.
Stakeholder Involvement	Methodology	Reliability/ Reproducibility	5. Targeted patient population. 6. Intended audience or users of the guideline.
4. Inclusion of all the relevant professional groups. 5. Patient preferences. 6. Target users. 7. Piloting among target users.	Reliable methods that integrate relevant research findings should be used to develop practice guidelines.	Practice guidelines are reliable and reproducible.	7. The principal preventive, diagnostic, or therapeutic options. 8. The health outcomes. 9. The method of external review. 10. An expiration date or date of scheduled review.
Rigor of Development	Clinical Expertise	Clinical Applicability	
8. Systematic methodology. 9. Selection criteria. 10. Methodology of formulating the recommendations. 11. Benefits and risks considered. 12. Link between the recommendations and the supporting evidence. 13. External review. 14. Procedure for update.	Appropriate clinical expertise should be used to develop practice guidelines.	Practice guidelines should be as inclusive of appropriately defined patient populations.	Standards of Evidence Identification and Summary
	Comprehensiveness	Clinical Flexibility	11. Method of identifying scientific evidence. 12. Time period from which evidence is reviewed. 13. Citations and references. 14. Method of data extraction. 15. Method for grading or classifying the scientific evidence. 16. Formal methods of combining evidence or expert opinion. 17. Specifics of benefits and harms of specific health practices. 18. Quantification of benefits and harms. 19. Effect on health care costs. 20. Quantification of costs.
	Practice guidelines should be as comprehensive and specific as possible.	Practice guidelines should identify the specifically known or generally expected exceptions to their recommendations.	
Clarity And Presentation	Current Information	Clarity	
15. Specific and unambiguous recommendations. 16. Clear presentation of options. 17. Key recommendations. 18. Supported with tools for application.	Practice guidelines should be based on current information.	Practice guidelines should use unambiguous language, define terms precisely, and use logical, easy-to-follow modes of presentation.	
	Dissemination	Multidisciplinary Process	
	Practice guidelines should be widely disseminated.	Practice guidelines should be developed by a process that includes participation by representatives of key affected groups.	Standards on the Formulation of Recommendations
Applicability		Scheduled Review	21. Value judgments used. 22. Patient preferences. 23. Specific recommendations and stated goals. 24. Grading of recommendations based on strength of the evidence. 25. Flexibility in the recommendations.
19. Potential organizational barriers discussed. 20. Potential cost implications considered. 21. Key review criteria presented.		Practice guidelines should include statements about when they should be reviewed.	
Editorial Independence		Documentation	
22. Editorial independence. 23. Report of conflicts of interest.		Description of the procedures followed in developing guidelines, the participants involved, the evidence used, the assumptions and rationales accepted, and the analytic methods employed.	

Table 11. *Sequential process for developing guidelines.*

<p><b>First steps</b></p> <p>1. <i>Establishing the process</i>—For example, prioritizing problems, selecting a panel, declaring conflicts of interest, and agreeing on group processes.</p> <p><b>Preparatory steps</b></p> <p>2. <i>Systematic review</i>—The first step is to identify and critically appraise or prepare systematic reviews of the best available evidence for all important outcomes.</p> <p>3. <i>Prepare evidence profile for important outcomes</i>—Profiles are needed for each subpopulation or risk group, based on the results of systematic reviews, and should include a quality assessment and a summary of findings.</p> <p><b>Grading quality of evidence and strength of recommendations</b></p> <p>4. <i>Quality of evidence for each outcome</i>—Judged on information summarized in the evidence profile.</p> <p>5. <i>Relative importance of outcomes</i>—Only important outcomes should be included in evidence profiles. The included outcomes should be classified as critical or important (but not critical) to a decision.</p> <p>6. <i>Overall quality of evidence</i>—The overall quality of evidence should be judged across outcomes based on the lowest quality of evidence for any of the critical outcomes.</p> <p>7. <i>Balance of benefits and harms</i>—The balance of benefits and harms should be classified as net benefits, trade-offs, uncertain trade-offs, or no net benefits based on the important health benefits and harms.</p> <p>8. <i>Balance of net benefits and costs</i>—Are incremental health benefits worth the costs? Because resources are always limited, it is important to consider costs (resource utilization) when making a recommendation.</p> <p>9. <i>Strength of recommendation</i>—Recommendations should be formulated to reflect their strength—that is, the extent to which one can be confident that adherence will do more good than harm.</p> <p><b>Subsequent steps</b></p> <p>10. <i>Implementation and evaluation</i>—For example, using effective implementation strategies that address barriers to change, evaluation of implementation, and keeping up to date.</p>
---

Adapted from Atkins D et al. Grading quality of evidence and strength of recommendations. *BMJ* 2004; 328:1490 (244).

compared: definitions, judgements, key components of the quality of evidence, summaries of evidence and findings, and extent of use. The advantages of the GRADE system are illustrated in Table 12.

Judgements about quality of evidence should be guided by a systematic review of available evidence. Reviewers should consider 4 key elements: study design, study quality, consistency, and directness. Study design refers to the basic study design, categorized into observational studies and randomized trials. Study quality refers to the detailed study methods and execution. Consistency refers to the similarity of estimates of effect across studies. Directness refers to the extent to which people, interventions, and outcome measures are similar to those of interest. The quality of evidence for each outcome can be determined after considering each of the 4 elements. Initially, the type

of study is considered and quality and limitations are looked into, which may alter the grading of evidence. The same rules should be applied to judgements about the quality of evidence for harms and benefits. Judgements about the quality of evidence for important studies can and should be made in the context of systematic reviews. Recommendations involve a trade-off between benefits and harms, making the trade-off inevitably involves placing, implicitly or explicitly, a relative value on each outcome.

The GRADE authors (244) suggest making explicit judgements about the balance between the main health benefits and harms before considering costs. In summary, recommendations should consider 4 main factors: the trade-offs, taking into account the estimated size of the effect for the main outcomes, the confidence limits around those estimates, and the rel-

Table 12. Comparison of *GRADE* and other systems.

Factor	Other systems	GRADE	Advantages of GRADE system *
Definitions	Implicit definitions of quality (level) of evidence and strength of recommendation	Explicit definitions	Makes clear what grades indicate and what should be considered in making these judgments
Judgements	Implicit judgments regarding which outcomes are important, quality of evidence for each important outcome, overall quality of evidence, balance between benefits and harms, and value of incremental benefits.	Sequential, explicit judgments	Clarifies each of these judgments and reduces risks of introducing errors or bias that can arise when they are made implicitly
Key components of quality of evidence	Not considered for each important outcome. Judgments about quality of evidence are often based on study design alone	Systematic and explicit consideration of study design, study quality, consistency, and directness of evidence in judgments about quality of evidence	Ensures these factors are considered appropriately
Other factors that can affect quality of evidence	Not explicitly taken into account	Explicit consideration of imprecise or sparse data, reporting bias, strength of association, evidence of a dose-response gradient, and plausible confounding	Ensures consideration of other factors
Overall quality of evidence	Implicitly based on the quality of evidence for benefits	Based on the lowest quality of evidence for any of the outcomes that are critical to making a decision	Reduces likelihood of mislabelling overall quality of evidence when evidence for a critical outcome is lacking
Relative importance of outcomes	Considered implicitly	Explicit judgments about which outcomes are critical, which ones are important but not critical, and which ones are unimportant and can be ignored	Ensures appropriate consideration of each outcome when grading overall quality of evidence and strength of recommendations
Balance between health benefits and harms	Not explicitly considered	Explicit consideration of trade-offs between important benefits and harms, the quality of evidence for these, translation of evidence into specific circumstances, and certainty of baseline risks	Clarifies and improves transparency of judgments on harms and benefits
Whether incremental health benefits are worth the costs	Not explicitly considered	Explicit consideration after first considering whether there are net health benefits	Ensures that judgments about value of net health benefits are transparent
Summaries of evidence and findings	Inconsistent presentation	Consistent GRADE evidence profiles, including quality assessment and summary of findings	Ensures that all panel members base their judgments on same information and that this information is available to others
Extent of use	Seldom used by more than one organization and little, if any, empirical evaluation	International collaboration across wide range of organizations in development and evaluation	Builds on previous experience to achieve a system that is more sensible, reliable, and widely applicable

\*Most other approaches do not include any of these advantages, although some may incorporate some of these advantages.

Source: Atkins D et al. Grading quality of evidence and strength of recommendations. *BMJ* 2004; 328:1490 (244).

ative value placed on each outcome; the quality of the evidence; translation of the evidence into practice in a specific setting, taking into consideration important factors that could be expected to modify the size of the expected effects, such as proximity to a hospital or availability of necessary expertise; and the uncertainty about baseline risk for the population of interest.

With any recommendation except for high, further research is required. With high grades of evidence further research is unlikely to change the confidence in the estimate effect. For moderate evidence, further research is very likely to have an important impact on low quality evidence, and any estimate of the effect is uncertain in very low quality evidence.



**ACKNOWLEDGMENTS**

The authors wish to thank the editorial board of *Pain Physician*, for review and criticism in improving the manuscript; Sekar Edem for his assistance in

search of literature; and Tonie M. Hatton and Diane E. Neihoff, transcriptionists (Pain Management Center of Paducah), for their assistance in preparation of this manuscript.

**REFERENCES**

- Sniderman AD, Furberg CD. Why guideline-making requires reform. *JAMA* 2009; 301:429-431.
- Eden J, Wheatley B, McNeil B, Sox H. Developing trusted clinical practice guidelines. In *Knowing What Works in Health Care: A Roadmap for the Nation*. National Academies Press, 2008, Washington, DC, pp 121-152.
- Field MJ, Lohr KN (eds). Committee to Advise the Public Health Service on Clinical Practice Guidelines, Institute of Medicine. *Clinical Practice Guidelines. Directions for a New Program*. National Academy Press, Washington, 1990.
- Sackett DL, Rosenberg WM, Gray JA, Haynes RB, Richardson WS. Evidence based medicine: What it is and what it isn't. *BMJ* 1996; 312:71-72.
- Timmermans S, Mauck A. The promises and pitfalls of evidence-based medicine. *Health Aff* 2005; 24:18-28.
- Sackett D, Richardson WS, Rosenberg W, Haynes B. *Evidence-based Medicine*. Churchill Livingstone, London, 1996.
- Eden J, Wheatley B, McNeil B, Sox H. *Knowing What Works in Health Care: A Roadmap for the Nation*. National Academies Press, Washington, DC, 2008.
- U.S. Department of Health and Human Services Agency for Healthcare Research and Quality (AHRQ) National Guideline Clearinghouse Guideline Index. [www.guideline.gov/browse/guideline\\_index.aspx](http://www.guideline.gov/browse/guideline_index.aspx)
- U.S. Department of Health and Human Services Agency for Healthcare Research and Quality (AHRQ). [www.ahrq.gov](http://www.ahrq.gov)
- American College of Occupational and Environmental Medicine. Low Back Disorders Chapter. In *Occupational Medicine Practice Guidelines: Evaluation and Management of Common Health Problems and Functional Recovery of Workers*, Second Edition. American College of Occupational and Environmental Medicine, Elk Grove Village, 2007.
- American College of Occupational and Environmental Medicine. Chronic Pain Chapter (revised 2008). In *Occupational Medicine Practice Guidelines: Evaluation and Management of Common Health Problems and Functional Recovery of Workers*, Second Edition. American College of Occupational and Environmental Medicine, Elk Grove Village, Epublished August 14, 2008.
- HAYES, Inc. Independent Health Technology Assessment Company. [www.hayesinc.com](http://www.hayesinc.com)
- Dennison PL. *Official Disability Guidelines*, 13th ed. Work Loss Data Institute, 2008.
- Phurrough S, Salive M, O'Connor D, Schafer J. *Decision Memo for Thermal Intradiscal Procedures*. 2008 [cited September 30, 2008]. [www.cms.hhs.gov/mcd/viewdecisionmemo.asp?from2=viweddecisionmemo.asp&id=215&](http://www.cms.hhs.gov/mcd/viewdecisionmemo.asp?from2=viweddecisionmemo.asp&id=215&)
- Phurrough S, Salive M, O'Connor D, Schafer J. *Proposed Coverage Decision Memorandum for Lumbar Artificial Disc Replacement*. [cited May 25, 2007]. [www1.cms.hhs.gov/mcd/viewdraftdecisionmemo.asp?from2=viewdraftdecisionmemo.asp&id=197&](http://www1.cms.hhs.gov/mcd/viewdraftdecisionmemo.asp?from2=viewdraftdecisionmemo.asp&id=197&)
- Centers for Medicare and Medicaid Services (CMS). Health Technology Assessment (HTA) Database: Percutaneous kyphoplasty for vertebral fractures caused by osteoporosis and malignancy. Blue Cross Blue Shield Association (BCBS), Chicago, 2005.
- Anthem Blue Cross Blue Shield (BCBS) Health and Medical Insurance: [www.anthem.com](http://www.anthem.com)
- Aetna Health and Medical Insurance: [www.aetna.com/index.htm](http://www.aetna.com/index.htm)
- United Healthcare Health and Medical Insurance: [www.uhc.com](http://www.uhc.com)
- Cigna Health and Medical Insurance: [www.cigna.com](http://www.cigna.com)
- Bluegrass Family Health and Medical Insurance: [www.bgfh.com](http://www.bgfh.com)
- Boswell MV, Trescot AM, Datta S, Schultz DM, Hansen HC, Abdi S, Sehgal N, Shah RV, Singh V, Benjamin RM, Patel VB, Buenaventura RM, Colson JD, Cordero HJ, Epter RS, Jasper JF, Dunbar EE, Atluri SL, Bowman RC, Deer TR, Swicegood JR, Staats PS, Smith HS, Burton AW, Kloth DS, Giordano J, Manchikanti L. Interventional techniques: Evidence-based practice guidelines in the management of chronic spinal pain. *Pain Physician* 2007; 10:7-111.
- Boswell MV, Shah RV, Everett CR, Sehgal N, McKenzie-Brown AM, Abdi S, Bowman RC, Deer TR, Datta S, Colson JD, Spillane WF, Smith HS, Lucas-Levin LF, Burton AW, Chopra P, Staats PS, Wasserman RA, Manchikanti L. Interventional techniques in the management of chronic spinal pain: Evidence-based practice guidelines. *Pain Physician* 2005; 8:1-47.
- Manchikanti L, Staats PS, Singh V, Schultz DM, Vilims BD, Jasper JF, Kloth DS, Trescot AM, Hansen HC, Falasca TD, Racz GB, Deer T, Burton AW, Helm S, Lou L, Bakhit CE, Dunbar EE, Atluri SL, Calodney AK, Hassenbusch S, Feler CA. Evidence-based practice guidelines for interventional techniques in the management of chronic spinal pain. *Pain Physician* 2003; 6:3-80.
- Manchikanti L, Singh V, Kloth DS, Slipman CW, Jasper JF, Trescot AM, Varley KG, Atluri SL, Giron C, Curran MJ, Rivera JJ, Baha A, Bakhit CE, Reuter M. Interventional techniques in the management of chronic pain: Part 2.0. *Pain Physician* 2001; 4:24-96.
- Manchikanti L, Singh V, Bakhit CE, Fellows B. Interventional techniques in the management of chronic pain. Part 1.0. *Pain Physician* 2000; 3:7-42.
- Manchikanti L, Heavner J, Racz GB, Mekhail NA, Schultz DM, Hansen HC, Singh V. Methods for evidence synthesis in interventional pain management. *Pain Physician* 2003; 6:89-111.
- Manchikanti L, Abdi S, Lucas LF. Evidence synthesis and development of guidelines in interventional pain management. *Pain Physician* 2005; 8:73-86.
- Manchikanti L, Boswell MV, Giordano J. Evidence-based interventional pain management: Principles, problems, po-

- tential, and applications. *Pain Physician* 2007; 10:329-356.
30. Manchikanti L. Evidence-based medicine, systematic reviews, and guidelines in interventional pain management: Part 1: Introduction and general considerations. *Pain Physician* 2008; 11:161-186.
  31. Manchikanti L, Hirsch JA, Smith HS. Evidence-based medicine, systematic reviews, and guidelines in interventional pain management: Part 2: Randomized controlled trials. *Pain Physician* 2008; 11:717-773.
  32. Manchikanti L, Smith HS, Hirsch JA. Evidence-based medicine, systematic reviews, and guidelines in interventional pain management: Part 3: Systematic reviews and meta-analysis of randomized trials. *Pain Physician* 2009; 12:35-72.
  33. Manchikanti L, Smith HS, Hirsch JA. Evidence-based medicine, systematic reviews, and guidelines in interventional pain management: Part 4: Observational studies. *Pain Physician* 2009; 12:73-108.
  34. Abdi S, Datta S, Trescot AM, Schultz DM, Adlaka R, Atluri SL, Smith HS, Manchikanti L. Epidural steroids in the management of chronic spinal pain: A systematic review. *Pain Physician* 2007; 10:185-212.
  35. Datta S, Everett CR, Trescot AM, Schultz DM, Adlaka R, Abdi S, Atluri SL, Smith HS, Shah RV. An updated systematic review of the diagnostic utility of selective nerve root blocks. *Pain Physician* 2007; 10:113-128.
  36. Trescot AM, Chopra P, Abdi S, Datta S, Schultz DM. Systematic review of the effectiveness and complications of adhesiolysis in the management of chronic spinal pain: An update. *Pain Physician* 2007; 10:129-146.
  37. Hansen HC, McKenzie-Brown AM, Cohen SP, Swicegood JR, Colson JD, Manchikanti L. Sacroiliac joint interventions: A systematic review. *Pain Physician* 2007; 10:165-184.
  38. Buenaventura RM, Shah RV, Patel V, Benyamin RM, Singh V. Systematic review of discography as a diagnostic test for spinal pain: An update. *Pain Physician* 2007; 10:147-164.
  39. Sehgal N, Dunbar EE, Shah RV, Colson JD. Systematic review of the diagnostic utility of facet (zygapophysial) joint injections in chronic spinal pain: An update. *Pain Physician* 2007; 10:213-228.
  40. Boswell MV, Colson JD, Sehgal N, Dunbar E, Epter R. A systematic review of therapeutic facet joint interventions in chronic spinal pain. *Pain Physician* 2007; 10:229-253.
  41. Atluri S, Datta S, Falco FJ, Lee M. Systematic review of diagnostic utility and therapeutic effectiveness of thoracic facet joint interventions. *Pain Physician* 2008; 11:611-629.
  42. Singh V, Manchikanti L, Shah RV, Dunbar EE, Glaser SE. Systematic review of thoracic discography as a diagnostic test for chronic spinal pain. *Pain Physician* 2008; 11:631-642.
  43. Patel VB, Manchikanti L, Singh, Schultz DM, Smith HS. Systematic review of intrathecal infusion systems for long-term management of chronic non-cancer pain. *Pain Physician* 2009; 12:345-360.
  44. Conn A, Buenaventura R, Datta S, Abdi S, Diwan S. Systematic review of caudal epidural injections in the management of chronic low back pain. *Pain Physician* 2009; 12:109-135.
  45. Parr AT, Diwan S, Abdi S. Lumbar interlaminar epidural injections in managing chronic low back and lower extremity pain: A systematic review. *Pain Physician* 2009; 12:163-188.
  46. Benyamin RM, Singh V, Parr AT, Conn A, Diwan S, Abdi S. Systematic review of the effectiveness of cervical epidurals in the management of chronic neck pain. *Pain Physician* 2009; 12:137-157.
  47. Buenaventura R, Datta S, Abdi S, Smith HS. Systematic review of therapeutic lumbar transforaminal epidural steroid injections. *Pain Physician* 2009; 12:233-251.
  48. Helm S, Hayek S, Benyamin RM, Manchikanti L. Systematic review of the effectiveness of thermal annular procedures in treating discogenic low back pain. *Pain Physician* 2009; 12:207-232.
  49. Manchikanti L, Dunbar EE, Wargo BW, Shah RV, Derby R, Cohen SP. Systematic review of cervical discography as a diagnostic test for chronic spinal pain. *Pain Physician* 2009; 12:305-321.
  50. Datta S, Lee M, Falco FJE, Bryce DA, Hayek SM. Systematic assessment of diagnostic accuracy and therapeutic utility of lumbar facet joint interventions. *Pain Physician* 2009; 12:437-460.
  51. Falco FJE, Erhart S, Wargo BW, Bryce DA, Atluri S, Datta S, Hayek SM. Systematic review of diagnostic utility and therapeutic effectiveness of cervical facet joint interventions. *Pain Physician* 2009; 12:323-344.
  52. Frey ME, Manchikanti L, Benyamin RM, Schultz DM, Smith HS. Spinal cord stimulation for patients with failed back surgery syndrome: A systematic review. *Pain Physician* 2009; 379-397.
  53. Epter RS, Helm S, Hayek SM, Benyamin RM, Smith HS, Abdi S. Systematic Review of percutaneous adhesiolysis and management of chronic low back pain in post lumbar surgery syndrome. *Pain Physician* 2009; 12:361-378.
  54. Smith HS, Chopra P, Patel VB, Frey ME, Rastogi R. Systematic review on the role of sedation in diagnostic spinal interventional techniques. *Pain Physician* 2009; 12:195-206.
  55. Rupert MP, Lee M, Manchikanti L, Datta S, Cohen SP. Evaluation of sacroiliac joint interventions: A systematic appraisal of the literature. *Pain Physician* 2009; 12:399-418.
  56. Manchikanti L, Glaser S, Wolfer L, Derby R, Cohen SP. Systematic review of lumbar discography as a diagnostic test for chronic low back pain. *Pain Physician* 2009; 12:541-559.
  57. Hayek SM, Helm S, Benyamin RM, Singh V, Bryce DA, Smith HS. Effectiveness of spinal endoscopic adhesiolysis in post lumbar surgery syndrome: A systematic review. *Pain Physician* 2009; 12:419-435.
  58. Crawford CC, Huynh MT, Kepple A, Jonas WB. Systematic Assessment of the Quality of Research Studies of Conventional and Alternative Treatment(s) of Primary Headache. *Pain Physician* 2009; 12:461-470.
  59. Trescot AM, Helm S, Hansen H, Benyamin RM, Adlaka R, Patel S, Manchikanti L. Opioids in the management of chronic non-cancer pain: An update of American Society of Interventional Pain Physicians' (ASIPP) guidelines. *Pain Physician* 2008; 11:S5-S62.
  60. Manchikanti L, Singh V, Derby R, Schultz DM, Benyamin RM, Prager JP, Hirsch JA. Reassessment of evidence synthesis of occupational medicine practice guidelines for interventional pain management. *Pain Physician* 2008; 11:393-482.
  61. Manchikanti L, Singh V, Derby R, Helm S, Trescot AM, Staats PS, Prager JP, Hirsch JA. Review of occupational medicine practice guidelines for interventional pain management and potential implications. *Pain Physician* 2008;

- 11:271-289.
62. Manchikanti L, Singh V, Helm S, Trescott AM, Hirsch JA. A critical appraisal of 2007 American College of Occupational and Environmental Medicine (ACOEM) practice guidelines for interventional pain management: An independent review utilizing AGREE, AMA, IOM, and other criteria. *Pain Physician* 2008; 11:291-310.
  63. Helm S. California workers' compensation system: Are occupational medicine practice guidelines sufficient for the treatment of chronic spinal pain or do they require supplementation by guidelines for interventional techniques? *Pain Physician* 2004; 7:229-238.
  64. The Cochrane Collaboration. *Evidence Health Care*. Available at: [www.cochrane.org/docs/ebm.htm](http://www.cochrane.org/docs/ebm.htm)
  65. Adminastar Federal Pain Management Policy. LCD Database ID Number L28529. Effective Date 01/01/2009.
  66. Wilson PR, Caplan RA, Connis RT, Gilbert HC, Grigsby EJ, Haddox JD, Harvey AM, Korevaar WC, Lubenow TR, Simon DL. Practice guidelines for chronic pain management. A report by the American Society of Anesthesiologists Task Force on Pain Management, Chronic Pain Section. *Anesthesiology* 1997; 86:995-1004.
  67. Sanders SH, Harden RN, Benson SE, Vicente PJ. Clinical practice guidelines for chronic non-malignant pain syndrome patients II: An evidence-based approach. *J Back Musc Rehabil* 1999; 13:47-58.
  68. Nelemans PJ, Debie RA, DeVet HC, Sturmans F. Injection therapy for subacute and chronic benign low back pain. *Spine* 2001; 26:501-515.
  69. Geurts JW, van Wijk RM, Stolker RJ, Groen GJ. Efficacy of radiofrequency procedures for the treatment of spinal pain: A systematic review of randomized clinical trials. *Reg Anesth Pain Med* 2001; 26:394-400.
  70. Manchikanti L, Singh V, Vilims BD, Hansen HC, Schultz DM, Kloth DS. Medial branch neurotomy in management of chronic spinal pain: Systematic review of the evidence. *Pain Physician* 2002; 5:405-418.
  71. Niemisto L, Kalso E, Malmivaara A, Seitsalo S, Hurri H. Cochrane Collaboration Back Review Group. Radiofrequency denervation for neck and back pain: A systematic review within the framework of the Cochrane collaboration back review group. *Spine* 2003; 28:1877-1888.
  72. van Tulder MW, Koes B, Malmivaara A. Outcome of non-invasive treatment modalities on back pain: An evidence-based review. *Eur Spine J* 2006; 15:S64-S81.
  73. Bogduk N, McGuirk B. *Medical Management of Acute and Chronic Low Back Pain. An Evidence-Based Approach: Pain Research and Clinical Management*. Elsevier Science BV, Amsterdam, 2002.
  74. Bogduk N, McGuirk B. *Management of Acute and Chronic Neck Pain. An Evidence-Based Approach*. Elsevier, Edinburgh, 2006.
  75. Taylor RS, Van Buyten JP, Buchser E. Spinal cord stimulation for chronic back and leg pain and failed back surgery syndrome: A systematic review and analysis of prognostic factors. *Spine* 2005; 30:152-160.
  76. Koes BW, Scholten RJ, Mens JMA, Bouter LM. Epidural steroid injections for low back pain and sciatica. An updated systematic review of randomized clinical trials. *Pain Digest* 1999; 9:241-247.
  77. Bigos SJ, Boyer OR, Braen GR, Brown K, Deyo R, Haldeman S, Hart JL, Johnson EW, Keller R, Kido D, Liang MH, Nelson RM, Nordin M, Owen BD, Pope MH, Schwartz RK, Stewart DH, Susman J, Triano JJ, Tripp LC, Turk DC, Watts C, Weinstein JN. Acute Low Back Problems in Adults. Clinical Practice Guideline No. 14, AHCPR Publication No. 95-0642. Rockville, Maryland, U.S.A., Agency for Health Care Policy and Research, Public Health Service, U.S., Department of Health and Human Services, December, 1994, pp 1-60.
  78. van Tulder MW, Koes BW, Bouter LM. Conservative treatment of acute and chronic nonspecific low back pain. A systematic review of randomized controlled trials of the most common interventions. *Spine* 1997; 22:2128-2156.
  79. Turner J, Loeser J, Deyo R, Sanders SB. Spinal cord stimulation for patients with failed back surgery syndrome or complex regional pain syndrome: A systematic review of effectiveness and complications. *Pain* 2004; 108:137-147.
  80. Airaksinen O, Brox JJ, Cedraschi C, Hildebrandt J, Kluber-Moffett J, Kovacs F, Mannion AF, Reis S, Staal JB, Ursin H, Zanoli G. Chapter 4: European guidelines for the management of chronic nonspecific low back pain. *Eur Spine J* 2006; 15:S192-S300.
  81. Mailis-Gagnon A, Furlan AD, Sandoval JA, Taylor R. Spinal cord stimulation for chronic pain. *Cochrane Database Syst Rev* 2004; 3:CD003783.
  82. Turner JA, Sears JM, Loeser JD. Programmable intrathecal opioid delivery systems for chronic non-malignant pain: A systematic review of effectiveness and complications. *Clin J Pain* 2007; 23:180-195.
  83. Turner JA, Loeser JD, Bell KG. Spinal cord stimulation for chronic low back pain. A systematic literature synthesis. *Neurosurgery* 1995; 37:1088-1096.
  84. Taylor RS. Spinal cord stimulation in complex regional pain syndrome and refractory neuropathic back and leg pain/failed back surgery syndrome: Results of a systematic review and meta-analysis. *J Pain Symptom Manage* 2006; 31:S13-S19.
  85. Taylor RS, Van Buyten JP, Buchser E. Spinal cord stimulation for complex regional pain syndrome: A systematic review of the clinical and cost-effectiveness literature and assessment of prognostic factors. *Eur J Pain* 2006; 10:91-101.
  86. Andersson GB, Mekhail NA, Block JE. Treatment of intractable discogenic low back pain. A systematic review of spinal fusion and intradiscal electrothermal therapy (IDET). *Pain Physician* 2006; 9:237-248.
  87. Appleby D, Andersson G, Totta M. Meta-analysis of the efficacy and safety of intradiscal electrothermal therapy (IDET). *Pain Med* 2006; 4:308-316.
  88. Gibson JN, Waddell G. Surgery for degenerative lumbar spondylosis: Updated Cochrane Review. *Spine* 2005; 30:2312-2320.
  89. Percutaneous Discectomy. Washington State Department of Labor and Industries, Office of Medical Director; February 24, 2004.
  90. Intradiscal Electrothermal Therapy (IDET). Technology Assessment Update. Washington State Department of Labor and Industries, Office of the Medical Director; September 30, 2003.
  91. Taylor RS, Taylor RJ, Fritzell P. Balloon kyphoplasty and vertebroplasty for vertebral compression fractures: A comparative systematic review of efficacy and safety. *Spine* 2006; 31:2747-2755.
  92. Hulme PA, Krebs J, Ferguson SJ, Ber-

- lemann U. Vertebroplasty and kyphoplasty: A systematic review of 69 clinical studies. *Spine* 2006; 31:1983-2001.
93. Staal JB, de Bie R, de Vet HC, Hildebrandt J, Nelemans P. Injection therapy for subacute and chronic low-back pain: An updated Cochrane review. *Spine* 2009; 34:49-59.
  94. Carragee EJ, Hurwitz EL, Cheng I, Carroll LJ, Nordin M, Guzman J, Peloso P, Holm LW, Côté P, Hogg-Johnson S, van der Velde G, Cassidy JD, Haldeman S, Bone and Joint Decade 2000-2010 Task Force on Neck Pain and Its Associated Disorders. Treatment of neck pain: Injections and surgical interventions: Results of the Bone and Joint Decade 2000-2010 Task Force on Neck Pain and Its Associated Disorders. *Spine* 2008; 33:S153-S169.
  95. Wolfer L, Derby R, Lee JE, Lee SH. Systematic review of lumbar provocation discography in asymptomatic subjects with a meta-analysis of false-positive rates. *Pain Physician* 2008; 11:513-538.
  96. Gill JB, Kuper M, Chin PC, Zhang Y, Schutt R Jr. Comparing pain reduction following kyphoplasty and vertebroplasty for osteoporotic vertebral compression fractures. *Pain Physician* 2007; 10:583-590.
  97. Jasper JF, Hayek SM. Implanted occipital nerve stimulators. *Pain Physician* 2008; 11:187-200.
  98. Haynes RB, Sackett RB, Gray JMA, Cook DC, Guyatt GH. Transferring evidence from research into practice, 1: The role of clinical care research evidence in clinical decisions. *ACPJ Club* 1996; 125: A14-A16.
  99. Guyatt G, Drummond R. Part 1. The basics: Using the medical literature. 1A. Introduction: The philosophy of evidence-based medicine. In *Users' Guides to the Medical Literature. A Manual for Evidence-Based Clinical Practice*. The Evidence-Based Medicine Working Group. AMA Press, 2002, pp 3-12.
  100. Napodano RJ. *Values in Medical Practice*. Human Sciences Press, New York, 1986.
  101. Hotopf M. The pragmatic randomized controlled trial. *Adv Psychiatr Treat* 2002; 8:326-333.
  102. Williams DD, Garner J. The case against 'the evidence': A different perspective on evidence-based medicine. *Br J Psychiatry* 2002; 180:8-12.
  103. Hotopf M, Lewis G, Normand C. Putting trials on trial: The costs and consequences of small trials in depression: A systematic review of methodology. *J Epidemiol Community Health* 1997; 51:354-358.
  104. Hotopf M, Churchill R, Lewis G. Pragmatic randomized controlled trials in psychiatry. *Br J Psychiatry* 1999; 175:217-223.
  105. Miles A, Charlton B, Bentley P, Polychronis A, Grey J, Price N. New perspectives in the evidence-based healthcare debate. *J Eval Clin Pract* 2000; 6:77-84.
  106. Steinbrook R. Guidance for guidelines. *N Engl J Med* 2007; 356:331-333.
  107. Manchikanti L, Singh V, Pampati V, Smith HS, Hirsch J. Analysis of growth of interventional techniques in managing chronic pain in the medicare population: A 10-year evaluation from 1997 to 2006. *Pain Physician* 2009; 12:9-34.
  108. Manchikanti L, Rivera JJ, Pampati V, Damron KS, Beyer CD, Barnhill RC. Effectiveness of clinical guidelines in interventional pain management. *Pain Physician* 2002; 5:127-132.
  109. Grimshaw JM, Russell IT. Effect of clinical guidelines on medical practice: A systematic review of rigorous evaluations. *Lancet* 1993; 342:1317-1322.
  110. Mitchell JB, Ballard DJ, Whisnant JP, Ammering CJ, Samsa GP, Matchar DB. What role do neurologists play in determining the costs and outcomes of stroke patients? *Stroke* 1996; 27:1937-1943.
  111. Wong JH, Findlay JM, Suarez-Almazor ME. Regional performance of carotid endarterectomy. Appropriateness, outcomes, and risk factors for complications. *Stroke* 1997; 28:891-898.
  112. Gill P, Dowell AC, Neal RD, Smith N, Heywood P, Wilson AE. Evidence based general practice: a retrospective study of interventions in one training practice. *BMJ* 1996; 30:819-821.
  113. Manchikanti L, Singh V, Datta S, Cohen SP, Hirsch JA. Comprehensive Review of Epidemiology, Scope, and Impact of Spinal Pain. *Pain Physician* 2009; 12: E35-E70.
  114. The National Uniform Claims Committee. Specialty Designation for Interventional Pain Management- 09.
  115. Medicare Payment Advisory Commission Report to the Congress. Paying for Interventional Pain Services in Ambulatory Settings. December 2001.
  116. Wright RW, Brand RA, Dunn W, Spindler KP. How to write a systematic review. *Clin Orthop Relat Res* 2007; 455:23-29.
  117. Pai M, McCulloch M, Gorman JD, Pai N, Enanoria W, Kennedy G, Tharyan P, Colford JM. Systematic reviews and meta-analyses: An illustrated, step-by-step guide. *Natl Med J India* 2004; 17:86-95.
  118. Egger M, Smith GD, Altman DG (eds). *Systematic Reviews in Health Care. Meta-Analysis in Context*. BMJ Publishing Group, London, 2001.
  119. World Health Organization. *International Clinical Trials Registry Platform (ICTRP)*. 2007. [www.who.int/entity/ictcp/en/](http://www.who.int/entity/ictcp/en/)
  120. Altman DG, Schulz KF, Moher D, Egger M, Davidoff F, Elbourne D, Gøtzsche PC, Lang T; CONSORT GROUP (Consolidated Standards of Reporting Trials). The revised CONSORT statement for reporting randomized trials: Explanation and elaboration. *Ann Intern Med* 2001; 134:663-694.
  121. Piaggio G, Elbourne DR, Altman DG, Pocock SJ, Evans SJ, CONSORT Group. Reporting of noninferiority and equivalence randomized trials: An extension of the CONSORT statement. *JAMA* 2006; 295:1152-1160.
  122. Zwarenstein M, Treweek S, Gagnier JJ, Altman DG, Tunis S, Haynes B, Oxman AD, Moher D; CONSORT group; Pragmatic Trials in Healthcare (Practihc) group. Improving the reporting of pragmatic trials: An extension of the CONSORT statement. *BMJ* 2008; 337: a2390.
  123. Peipert JF, Phipps MG. Observational studies. *Clin Obstet Gynecol* 1998; 41:235-244.
  124. Stroup DF, Berlin JA, Morton SC, Olkin I, Williamson GD, Rennie D, Moher D, Becker BJ, Sipe TA, Thacker SB. Meta-analysis of observational studies in epidemiology: A proposal for reporting. Meta-analysis of Observational Studies in Epidemiology (MOOSE) group. *JAMA* 2000; 283:2008-2012.
  125. Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig LM, Lijmer JG, Moher D, Rennie D, de Vet HC; STARD Group. Towards complete and accurate reporting of studies of diagnostic accuracy: The STARD initiative. *Fam Pract* 2004; 21:4-10.
  126. Wu JS, Wong RK, Lloyd NS, Johnston M, Bezjak A, Whelan T; Supportive Care Guidelines Group of Cancer Care On-

- tario. Radiotherapy fractionation for the palliation of uncomplicated painful bone metastases - An evidence-based practice guideline. *BMC Cancer* 2004; 4:71.
127. Manchikanti L. Health care reform in the United States: Radical surgery needed now more than ever. *Pain Physician* 2008; 11:13-42.
  128. Letter to Honorable Representatives Gordon H. Smith and John D. Rockefeller IV from A. Bruce Steinwald, Director, Health Care, United States Government Accountability Office. RE: Medicare Trends in Fees, Utilization, and Expenditures for Imaging Services before and after Implementation of the Deficit Reduction Act of 2005. September 26, 2008.
  129. 2008 Annual Report of the Boards of Trustees of the Federal Hospital Insurance and Federal Supplementary Medical Insurance Trust Funds (Washington, D.C. March 25, 2008).
  130. Straus SE, Glasziou P, Richardson WS, Haynes RB. *Evidence-Based Medicine: How to Practice and Teach EBM*. 3rd ed. Churchill Livingstone, London, UK, 2005.
  131. Wennberg J, Gittelsohn A. Small area variation in health care delivery. *Sci Am* 1973; 182:1102-1108.
  132. Schuster MA, McGlynn EA, Brook RH. How good is the quality of health care in the United States? *Milbank Q* 1998; 76:517-563.
  133. Manchikanti L. Medicare in interventional pain management: A critical analysis. *Pain Physician* 2006; 9:171-198.
  134. Manchikanti L, Boswell MV. Interventional techniques in ambulatory surgical centers: A look at the new payment system. *Pain Physician* 2007; 10:627-650.
  135. Manchikanti L, Giordano J. Physician payment 2008 for interventionalists: Current state of health care policy. *Pain Physician* 2007; 10:607-626.
  136. Friedly J, Chan L, Deyo R. Increases in lumbosacral injections in the Medicare population: 1994 to 2001. *Spine* 2007; 32:1754-1760.
  137. US Department of Health and Human Services. Office of Inspector General (OIG). Medicare Payments for Facet Joint Injection Services (OEI-05-07-00200). September 2008. [www.oig.hhs.gov/oei/reports/oei-05-07-00200.pdf](http://www.oig.hhs.gov/oei/reports/oei-05-07-00200.pdf)
  138. Manchikanti L, Hirsch JA. Issues in health care: Interventional pain management at the crossroads. Health policy update. *Pain Physician* 2007; 10:261-284.
  139. Specialty Utilization data files from CMS: [www.cms.hhs.gov](http://www.cms.hhs.gov)
  140. Eddy DM, Billings J. The quality of medical evidence: Implications for quality of care. *Health Aff (Millwood)* 1988; 7:19-32.
  141. McNeil BJ. Shattuck lecture—hidden barriers to improvement in the quality of care. *N Engl J Med* 2001; 345:1612-1620.
  142. Evidence-based Practice Centers Partner's Guide. Prepared for Agency for Healthcare Research and Quality by the EPC Coordinating Center. January 2005. [www.ahrq.gov/Clinic/epcpartner/epcpartner.pdf](http://www.ahrq.gov/Clinic/epcpartner/epcpartner.pdf)
  143. Kohn LT, Corrigan JM, Donaldson MS, editors. To Err Is Human: Building a Safer Health System. A Report of the Committee on Quality of Health Care in America. Institute of Medicine. National Academy Press, Washington, DC, 2000.
  144. Institute of Medicine, Committee on Quality of Health Care in America. Crossing the Quality Chasm: A New Health System for the 21st Century. National Academy Press, Washington, DC, 2001.
  145. Haynes RB, Devereaux PJ, Guyatt GH. Clinical expertise in the era of EBM and patient choice. *ACP J Club* 2002; 136: A11-A14.
  146. Manchikanti L. Interventional pain management: Past, present, and future. The Prithvi Raj lecture: Presented at the 4th World Congress-World Institute of Pain, Budapest, 2007. *Pain Pract* 2007; 7:357-371.
  147. Manchikanti L, Boswell MV, Raj PP, Racz GB. Evolution of interventional pain management. *Pain Physician* 2003; 6:485-494.
  148. 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.
  149. Cates JR, Young DN, Bowerman DS, Porter RC. An independent AGREE evaluation of the Occupational Medicine Practice Guidelines. *Spine J* 2006; 6:72-77.
  150. Moore RA, Derry S, Phillips CJ, McQuay HJ. Nonsteroidal anti-inflammatory drugs (NSAIDs), cyclooxygenase-2 selective inhibitors (coxibs) and gastrointestinal harm: Review of clinical trials and clinical practice. *BMC Musculoskeletal Disord* 2006; 7:79.
  151. Fuhrer MJ. Overview of clinical trials in medical rehabilitation: Impetuses, challenges, and needed future directions. *Am J Phys Med Rehabil* 2003; 82: S8-15.
  152. Liddle SD, Baxter GD, Gracey JH. Exercise and chronic low back pain: What works? *Pain* 2004; 107:176-190.
  153. Murphy AY, van Teijlingen ER, Gobbi MO. Inconsistent grading of evidence across countries: A review of low back pain guidelines. *J Manipulative Physiol Ther* 2006; 29:576-581.
  154. American Society of Interventional Pain Physicians - Medical Specialty Society. Interventional techniques in the management of chronic spinal pain: Evidence-based practice guidelines. National Guideline Clearinghouse; 2003 (revised 2007), 105 pages, NGC:005510.
  155. American Society of Interventional Pain Physicians - Medical Specialty Society. Opioid guidelines in the management of chronic non-cancer pain. National Guideline Clearinghouse; 2006 (revised 2008), 58 pages, NGC:006445.
  156. Druss BG, Marcus SC. Growth and decentralization of the medical literature: Implications for evidence-based medicine. *J Med Libr Assoc* 2005; 93:499-501.
  157. Kelly JT, Toepp MC. Practice parameters: Development, evaluation, dissemination, and implementation. *QRB Qual Rev Bull* 1992; 18:405-409.
  158. American Medical Association. Office of Quality Assurance and Health Care Organizations: Directory of Practice Parameters. Chicago, 1992.
  159. Woolf SH. Practice guidelines: A new reality in medicine, I. Recent developments. *Arch Intern Med* 1990; 150:1811-1818.
  160. Woolf SH. Practice guidelines: A new reality. medicine, II. Methods of developing guidelines. *Arch Intern Med* 1992; 152:946-952.
  161. Leape LL. Practice guidelines and standards: An overview. *QRB Qual Rev Bull* 1990; 16:42-49.
  162. American Medical Association Specialty Society Practice Parameters Partner-

- ship and Practice Parameters Forum. *Attributes to Guide the Development of Practice Parameters*. Office of Quality Assurance and Medical Review of the American Medical Association, Chicago, 1994.
163. Carter A. Clinical practice guidelines. *Can Med Assoc J* 1992; 147:1649-1650.
  164. Snow J. On the fatal cases of the inhalation of chloroform. *Edinburgh Med Surg J* 1849; 72:75.
  165. Tinker JH, Dull DL, Caplan RA. Role of monitoring devices in the prevention of anesthetic mishaps: A closed claims analysis. *Anesthesiology* 1989; 71:541-546.
  166. Shomaker TS. Practice policies in anesthesia: A foretaste of practice in the twenty-first century. *Anesth Analg* 1995; 80:388-403.
  167. Eichhorn JH, Cooper JB, Cullen DJ, Maier WR, Philip JH, Seeman RG. Standards for patient monitoring during anesthesia at Harvard Medical School. *JAMA* 1986; 256:1017-1020.
  168. American Academy of Pediatrics. *Report of the Committee on Immunization Procedures of the American Academy of Pediatrics*. American Academy of Pediatrics, Evanston, 1938.
  169. Nunn CM. Pathways, guidelines, and cookbook medicine: Are we all becoming Betty Crocker? *JCOM* 1997; 4:17-24.
  170. Eden J, Wheatley B, McNeil B, Sox H. Building a foundation for knowing what works in health care. In *Knowing What Works in Health Care: A Roadmap for the Nation*. National Academies Press, Washington, DC, 2008, pp 153-178.
  171. Eccles M, Mason J. How to develop cost conscious guidelines. *Health Technol Assess* 2001; 5:1-69.
  172. West S, King V, Carey T, Lohr K, McKay N, Sutton S, Lux L. *Systems to Rate the Strength of Scientific Evidence*. Evidence Report/Technology Assessment No. 47 University of North Carolina: Agency for Healthcare Research and Quality. AHRQ Publication No. 02-E016; April 2002.
  173. Cates JR, Young DN, Guerriero DJ, Jahn WT, Armine JP, Korbett AB, Bowerman DS, Porter RC, Sandman TD, King RA. Evaluating the quality of clinical practice guidelines. *J Manipulative Physiol Ther* 2001; 24:170-176.
  174. van Tulder MW, Tuut M, Pennick V, Bombardier C, Assendelft W. Quality of primary care guidelines for acute low back pain. *Spine* 2004; 29:E357-E362.
  175. Oxman AD, Cook DJ, Guyatt GH. Users' guide to the medical literature. VI. How to use an overview. Evidence-based Medicine Working Group. *JAMA* 1994; 272:1367-1371.
  176. Oxman AD, Schünemann 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.
  177. Oxman AD. Systematic reviews: Checklists for review articles. *BMJ* 1994; 309:648-651.
  178. Hayward R, Wilson M, Tunis S, Bass E, Guyatt G. Users' guides to the medical literature. VIII. How to use clinical practice guidelines. *JAMA* 1995; 274:570-574.
  179. 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; 281:1900-1905.
  180. Shekelle P, Woolf S, Grimshaw J. Clinical guidelines: Developing guidelines. *BMJ* 1999; 318:593-596.
  181. Connis R, Nickinovich D, Caplan R, Arens J. The development of evidence-based clinical practice guidelines. Integrating medical science and practice. *Int J Technol Assess Health Care* 2000; 16:1003-1012.
  182. Silagy C, Rubin G, Henderson-Smart D, Gross P. A guide to the development, implementation and evaluation of clinical practice guidelines. National Health and Medical Research Council, Canberra, Commonwealth of Australia, 1998, pp 1-79.
  183. O'Brien P, Silagy C, McCallum J, O'Connell, D, Glasziou P, Hill A. How to use the evidence: Assessment and application of scientific evidence. National Health and Medical Research Council, Canberra, Commonwealth of Australia, 2000, pp 1-84.
  184. Cluzeau FA, Burgers JS, Brouwers M, Grol R, Makela M, Littlejohns P, Grimshaw J, Hunt C, Asua J, Bataillard A, Browman G, Burnand B, Durieux P, Fervers B, Grilli R, Hanna S, Have P, Jovell A, Klazinga N, Kristensen F, Madsen PB, Miller J, Ollenschlaeger G, Qureshi S, Rico-Iturrioz R, Vader J-P, Zaat J. 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.
  185. Thomason M, Cluzeau F, Littlejohns P, Ollenschlaeger G, Grilli R, Rico-Iturrioz R, Grol R, Burgers J, Kristensen FB, Palmhoj-Nielsen C, Fervers B, Jovell A, Moran S, Grimshaw J, Burnand B, Klazinga N, Ten Have P, Durieux P, Miller J, Feder G, Makela M. Guideline development in Europe: An international comparison. *Int J Technol Assess Health Care* 2000; 16:1039-1049.
  186. AGREE Collaboration (Appraisal of Guidelines, Research, and Evaluation in Europe [AGREE] Collaborative Group). Appraisal of Guidelines for Research & Evaluation (AGREE) Instrument, September 2001.
  187. Hasenfeld R, Shekelle PG. Is the methodological quality of guidelines declining in the US? Comparison of the quality of US Agency for Health Care Policy and research (AHCPR) guidelines with those published subsequently. *Qual Saf Health Care* 2003; 12:428-434.
  188. Grol R, Cluzeau FA, Burgers JS. Clinical practice guidelines: Towards better quality guidelines and increased international collaboration. *Br J Cancer* 2003; 89:S4-S8.
  189. Burgers JS, Fervers B, Haugh M, Brouwers M, Browman G, Philip T, Cluzeau FA. 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.
  190. AGREE Collaboration (Appraisal of Guidelines, Research, and Evaluation in Europe [AGREE] Collaborative Group). Appraisal of Guidelines for Research & Evaluation (AGREE) Instrument Training Manual January 2003.
  191. Nuckols TK, Winn BO, Lim Y-W, Shaw RN, Mattke S, Wickizer T, Harber P, Wallace P, Asch SM, MacLean C, Garland RH. Evaluating medical treatment guideline sets for injured workers in California. RAND Corporation. Institute for Civil Justice and RAND Health. www.rand.org/pubs/monographs/2005/RAND\_MG400.pdf
  192. Staal JB, Hlobil H, van Tulder MW, Waddell G, Burton AK, Koes BW, van Mechelen W. Occupational health guidelines for the management of low back pain: An international comparison. *Occup Environ Med* 2003; 60:618-626.
  193. Peloso PMJ, Gross A, Haines T, Trinh K, Goldsmith CH, Burnie SJ, Cervical Overview Group. Medicinal and injec-

- tion therapies for mechanical neck disorders. *Cochrane Database Syst Rev* 2007; 3:CD000319.
194. Cook DJ, Greengold NL, Ellrodt AG, Weingarten SR. The relation between systematic reviews and practice guidelines. *Ann Intern Med* 1997; 127:210-216.
  195. Schünemann HJ, Fretheim A, Oxman AD, WHO Advisory Committee on Health Research. Improving the use of research evidence in guideline development: 1. Guidelines for guidelines. *Health Res Policy Syst* 2006; 4:13.
  196. American Medical Association, Office of Quality Assurance. *Attributes to Guide the Development and Evaluation of Practice Parameters*. American Medical Association, Chicago, 1990.
  197. Canadian Medical Association. *Quality of Care Program: The Guidelines for Canadian Clinical Practice Guidelines*. Canadian Medical Association, Ottawa, Ontario, 1993.
  198. Woolf SH. *Manual for Clinical Practice Guideline Development*. Rockville, MD. (Agency for Healthcare Policy and Research, AHCPR publication 91-0007, 1991.)
  199. Eden J, Wheatley B, McNeil B, Sox H. An imperative for change. In *Knowing What Works in Health Care: A Roadmap for the Nation*. National Academies Press, Washington, DC, 2008, pp 33-56.
  200. Grilli R, Magrini N, Penna A, Mura G, Liberati A. Practice guidelines developed by specialty societies: The need for critical appraisal. *Lancet* 2000; 355:103-106.
  201. Gonzalez EG, Materson RS. The guidelines, the controversy, the book. In Gonzalez ER, Materson RS (eds). *The Nonsurgical Management of Acute Low Back Pain*. Demos Vermande, New York, 1997, pp 1-4.
  202. LaDou J, Teitelbaum DT, Egilman DS, Frank AL, Kramer SN, Huff J. American College of Occupational and Environmental Medicine (ACOEM): A professional association in service to industry. *Int J Occup Environ Health* 2007; 13:404-426.
  203. Egilman DS. Suppression bias at the *Journal of Occupational and Environmental Medicine*. *Int J Occup Environ Health* 2005; 11:202-204.
  204. Lenzer J. Alteplase for stroke: Money and optimistic claims buttress the "brain attack" campaign. *BMJ* 2002; 324:723-726.
  205. Lenzer J. US consumer body calls for review of cholesterol guidelines. *BMJ* 2004; 329:759.
  206. DeAngelis CD, Fontanarosa PB. Impugning the integrity of medical science: The adverse effects of industry influence. *JAMA* 2008; 299:1833-1835.
  207. Kassirer JP. *On the Take: How Medicine's Complicity With Big Business Can Endanger Your Health*. Oxford University Press, New York, 2005, pp 97-99.
  208. Choudhry NK, Stelfox HT, Detsky AS. Relationships between authors of clinical practice guidelines and the pharmaceutical industry. *JAMA* 2002; 287:612-617.
  209. Psaty BM, Furberg C. British guidelines on managing hypertension provide evidence, progress, and an occasional missed opportunity. *BMJ* 1999; 319:589-590.
  210. Abdi S, Lucas LF, Datta S. Role of epidural steroids in the management of chronic spinal pain: A systematic review of effectiveness and complications. *Pain Physician* 2005; 8:127-143.
  211. Boswell MV, Colson JD, Spillane WF. Therapeutic facet joint interventions: A systematic review of their role in chronic spinal pain management and complications. *Pain Physician* 2005; 8:101-114.
  212. Chopra P, Smith HS, Deer TR, Bowman RC. Systematic review of adhesiolysis in managing chronic low back pain. *Pain Physician* 2005; 8:87-100.
  213. Shah RV, Everett CR, McKenzie-Brown AM, Sehgal N. Discography as a diagnostic test for spinal pain: A systematic and narrative review. *Pain Physician* 2005; 8:187-209.
  214. Everett CR, Shah R, Sehgal N, McKenzie-Brown AM. A systematic review of diagnostic utility of selective nerve root blocks. *Pain Physician* 2005; 8:225-233.
  215. Sehgal N, Shah RV, McKenzie-Brown AM, Everett CR. Diagnostic utility of facet (zygapophysial) joint injections in chronic spinal pain: A systematic review of evidence. *Pain Physician* 2005; 8:211-224.
  216. McKenzie-Brown A, Shah RV, Sehgal N, Everett CR. A systematic review of sacroiliac joint interventions. *Pain Physician* 2005; 8:115-125.
  217. Jenicek M. How do we see medicine, health and disease? A basic set of rules and fundamental paradigms. In *Foundations of Evidence-Based Medicine*. The Parthenon Publishing Group, Inc., New York, 2005, pp, 3-13.
  218. Hatala R, Guyatt G. Evaluating the teaching of evidence-based medicine. *JAMA* 2002; 288:1110-1112.
  219. Guyatt G, Drummond R. Part 1. The basics: Using the medical literature. 1C2. Diagnostic tests. In *Users' Guides to the Medical Literature: A Manual for Evidence-Based Clinical Practice*. American Medical Association, Chicago, 2002, pp 121-140.
  220. Glasziou P, Vandenbroucke JP, Chalmers I. Assessing the quality of research. *BMJ* 2004; 328:39-41.
  221. Black N. Why we need observational studies to evaluate the effectiveness of health care. *BMJ* 1996; 312:1215-1218.
  222. Funai EF, Rosenbush EJ, Lee MJ, Del Priore G. Distribution of study designs in four major US journals of obstetrics and gynecology. *Gynecol Obstet Invest* 2001; 51:8-11.
  223. Scales CD Jr., Norris RD, Peterson BL, Preeminger GM, Dahm P. Clinical research and statistical methods in the urology literature. *J Urol* 2005; 174:1374-1379.
  224. Greene T. Are observational studies "just as effective" as randomized clinical trials? *Blood Purif* 2000; 18:317-322.
  225. Papanikolaou PN, Christidi GD, Ioannidis JP. Comparison of evidence on harms of medical interventions in randomized and nonrandomized studies. *CMAJ* 2006; 174:635-641.
  226. Hempel CG. *Philosophy of Natural Science*. Prentice-Hall, Englewood Cliffs, 1966.
  227. Popper KR. *The Logic of Scientific Discovery*. Cambridge University Press, London, 1958.
  228. Burgers JS, van Everdingen JJ. Beyond the evidence in clinical guidelines. *Lancet* 2004; 364:392-393.
  229. Atkins D, Siegel J, Slutsky J. Making policy when the evidence is in dispute. *Health Aff (Millwood)* 2005; 24:102-113.
  230. Ricci S, Celani MG, Righetti E. Development of clinical guidelines: Methodological and practical issues. *Neurol Sci* 2006; 27:S228-S230.
  231. Perlin JB, Kupersmith J. Information technology and the inferential gap. *Health Aff* 2007; 26:w192-w194.
  232. Stewart WF, Shah NR, Selna MJ, Paulus

- RA, Walker JM. Bridging the inferential gap: The electronic health record and clinical evidence. *Health Aff* 2007; 26: w181-w191.
233. Berg AO, Allan JD. Introducing the third U.S. Preventive Services Task Force. *Am J Prev Med* 2001; 20:S3-S4.
234. Eden J, Wheatley B, McNeil B, Sox H. Systematic reviews: The central link between evidence and clinical decision making. In *Knowing What Works in Health Care: A Roadmap for the Nation*. National Academies Press, Washington, DC, 2008, pp 81-120.
235. Chou R, Helfand M. Challenges in systematic reviews that assess treatment harms. *Ann Intern Med* 2005; 142:1090-1099.
236. Tatsioni A, Zarin DA, Aronson N, Samson DJ, Flamm CR, Schmid C, Lau J. Challenges in systematic reviews of diagnostic technologies. *Ann Intern Med* 2005; 142:1048-1055.
237. Norris SL, Atkins D. Challenges in using nonrandomized studies in systematic reviews of treatment interventions. *Ann Intern Med* 2005; 142:1112-1119.
238. Whiting P, Rutjes A, Reitsma J, Bossuyt P, Kleijnen J. The Development of QUADAS: A tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. *BMC Med Res Methodol* 2003; 3:25.
239. Tunis S. Improving evidence for health care decisions. Presentation to IOM Staff, April 28, 2006. Washington, DC.
240. Neumann PJ, Divi N, Beinfeld MT, Levine BS, Keenan PS, Halpern EF, Gazelle GS. Medicare's national coverage decisions, 1999-2003: Quality of evidence and review times. *Health Aff* 2005; 24:243-254.
241. Uhlig K, Balk EM, Lau J. Grading evidence-based guidelines--What are the issues? *Am J Kidney Dis* 2008; 52:211-215.
242. Guyatt G, Vist G, Falck-Ytter Y, Kunz R, Magrini N, Schunemann H. An emerging consensus on grading recommendations? *Evid Based Med* 2006; 11:2-4.
243. Guyatt G. An emerging consensus on grading recommendations? *Chin J Evid-based Med* 2007; 7:1-4.
244. Atkins D, Best D, Briss PA, Eccles M, Falck-Ytter Y, Flottorp S, Guyatt GH, Harbour RT, Haugh MC, Henry D, Hill S, Jaeschke R, Leng G, Liberati A, Magrini N, Mason J, Middleton P, Mrukowicz J, O'Connell D, Oxman AD, Phillips B, Schünemann HJ, Edejer TT, Varonen H, Vist GE, Williams JW Jr, Zaza S; GRADE Working Group. Grading quality of evidence and strength of recommendations. *BMJ* 2004; 328:1490.
245. 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. *Am Fam Physician* 2004; 69:548-556.
246. Schünemann HJ, Best D, Vist G, Oxman AD; GRADE Working Group. Letters, numbers, symbols and words: How to communicate grades of evidence and recommendations. *CMAJ* 2003; 169:677-680.
247. Atkins D, Eccles M, Flottorp S, Guyatt GH, Henry D, Hill S, Liberati A, O'Connell D, Oxman AD, Phillips B, Schünemann H, Edejer TT, Vist GE, Williams JW Jr; GRADE Working Group. Systems for grading the quality of evidence and the strength of recommendations I: Critical appraisal of existing approaches. *BMC Health Serv Res* 2004; 4:38.
248. Atkins D, Briss PA, Eccles M, Flottorp S, Guyatt GH, Harbour RT, Hill S, Jaeschke R, Liberati A, Magrini N, Mason J, O'Connell D, Oxman AD, Phillips B, Schünemann H, Edejer TT, Vist GE, Williams JW Jr; GRADE Working Group. 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.
249. Schünemann HJ, Fretheim A, Oxman AD. Improving the use of research evidence in guideline development: 9. Grading evidence and recommendations. *Health Res Policy Sys* 2006; 4:21.
250. Schünemann HJ, Oxman AD, Brozek J, Glasziou P, Jaeschke R, Vist GE, Williams JW Jr, Kunz R, Craig J, Montori VM, Bossuyt P, Guyatt GH; GRADE Working Group. Grading quality of evidence and strength of recommendations for diagnostic tests and strategies. *BMJ* 2008; 336:1106-1110.
251. Guyatt GH, Oxman AD, Kunz R, Falck-Ytter Y, Vist GE, Liberati A, Schünemann HJ; GRADE Working Group. Going from evidence to recommendations. *BMJ* 2008; 336:1049-1051.
252. Guyatt GH, Oxman AD, Kunz R, Vist GE, Falck-Ytter Y, Schünemann HJ; GRADE Working Group. What is "quality of evidence" and why is it important to clinicians? *BMJ* 2008; 336:995-998.
253. Guyatt GH, Oxman AD, Vist GE, Kunz R, Falck-Ytter Y, Alonso-Coello P, Schünemann HJ; GRADE Working Group. An emerging consensus on rating quality of evidence and strength of recommendations. *BMJ* 2008; 336:924-926.
254. Oxman AD, Glasziou P, Williams JW Jr. What should clinicians do when faced with conflicting recommendations? *BMJ* 2008; 337:a2530.
255. 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.
256. Kunz R. What's in the black box? *Chest* 2006; 129:7-10.
257. Council of Europe. Developing a methodology for drawing up guidelines on best medical practice: Committee of Ministers of the Council of Europe. [www.coe.int/T/E/Social\\_Cohesion/Health/Recommendations/](http://www.coe.int/T/E/Social_Cohesion/Health/Recommendations/). Accessed October 25, 2005.
258. National Health and Medical Research Council. How to use the evidence: Assessment and application of scientific evidence. *Commonwealth of Australia*, 2000.
259. Phillips B, Ball C, Sackett D, Badenoch D, Straus S, Hayes B, Dawes M. *Levels of Evidence and Grades of Recommendations*. Last Revised May 2001. Centre for Evidence-Based Medicine. [www.cebm.net/levels\\_of\\_evidence.asp](http://www.cebm.net/levels_of_evidence.asp)
260. Guyatt G, Drummond R. *Users' Guide to the Medical Literature*. AMA Press, Chicago, 2002, pp 55-154.
261. Higgins JPT, Green S (eds.). *Cochrane Handbook for Systematic Reviews of Interventions* Version 5.0.1 [updated September 2008]. The Cochrane Collaboration, 2008. Available from [www.cochrane-handbook.org](http://www.cochrane-handbook.org)
262. Evidence-Based Medicine Working Group. Evidence-based medicine. A new approach to teaching the practice of medicine. *JAMA* 1992; 268:2420-2425.
263. Slawson DC, Shaughnessy AF, Bennett JH. Becoming a medical information master: Feeling good about not knowing everything. *J Fam Pract* 1994; 38:505-513.
264. Shaughnessy AF, Slawson DC, Bennett JH. Becoming an information master: A guidebook to the medical information



- jungle. *J Fam Pract* 1994; 39:489-499.
265. Siwek J, Gourlay ML, Slawson DC, Shaughnessy AF. How to write an evidence-based clinical review article. *Am Fam Physician* 2002; 65:251-258.
266. Guyatt GH, Schünemann H, Cook D, Pauker S, Sinclair J, Bucher H, Jaeschke R. Grades of recommendation for anti-thrombotic agents. *Chest* 2001; 119:3S-7S.
267. Harbour R, Miller J. A new system for grading recommendations in evidence based guidelines. *BMJ* 2001; 323:334-336.
268. Atkins D, Best D, Shapiro EN. The third U.S. Preventive Services Task Force: Background, methods and first recommendations. *Am J Preventive Medicine* 2001; 20:1-108.
269. Lepor H, Williford WO, Barry MJ, Brawer MK, Dixon CM, Gormley G, Haakenson C, Machi M, Narayan P, Padley RJ. The efficacy of terazosin, finasteride, or both in benign prostatic hyperplasia. Veterans Affairs Cooperative Studies Benign Prostatic Hyperplasia Study Group. *N Engl J Med* 1996; 335:533-539.
270. Briss PA, Zaza S, Pappaioanou M, Fielding J, Wright-De Agüero L, Truman BI, Hopkins DP, Mullen PD, Thompson RS, Woolf SH, Carande-Kulis VG, Anderson L, Hinman AR, McQueen DV, Teutsch SM, Harris JR. Developing an evidence-based guide to community preventive services—methods. The Task Force on Community Preventive Services. *Am J Prev Med* 2000; 18:S35-S43.

