Pain Physician. 2005;8:73-86, ISSN 1533-3159

An Invited Review

Evidence Synthesis and Development of Guidelines in Interventional Pain Management

Laxmaiah Manchikanti, MD, Salahadin Abdi, MD, PhD, and Linda F. Lucas, MD

Background: The past decade has been marked by unprecedented interest in evidence-based medicine and various types of avenues that can provide valid and reliable information about healthcare, including clinical practice guidelines. Thus, healthcare decisions are purportedly and increasingly being made on research-based evidence rather than on expert opinion or clinical experience alone. The methods describing evidence synthesis or development of guidelines in interventional pain management are scarce.

Objective: To describe the methods of evidence synthesis and development of guidelines in interventional pain management in practical terms.

Description: This review sought to provide an understanding of evidence-based medicine, the importance and synthesis of clinical practice guidelines, and the variety of issues related to rating the quality of individual articles, understanding narrative and systematic reviews, grading the strength of the body of evidence, searching for evidence, and finally, the analytical preparation of guidelines, with a focus on interventional pain management.

Conclusion: Evidence synthesis and development of guidelines in interventional pain management is a complex and difficult task. It is emphasized that practice guidelines are not intended as standards or absolute requirements. Consequently, practice guidelines may be adapted, modified, or rejected, according to the clinical needs and constraints of each practitioner and patient.

Keywords: Interventional pain management, evidence synthesis, clinical guidelines, evidence-based medicine, pragmatic or practical clinical trials, randomized trials, observational studies

The need for careful scientific evaluation of clinical practice became a prominent focus during the second half of the twentieth century (1). Tunis et al (1), affiliated with Centers for Medicare and Medicaid Services and the Agency for Health Care Research and Quality, described that the demonstration of pervasive and persistent unexplained variations in clinical practice (2), and high rates of inappropriate care (3), combined with increased expenditures, have fueled a steadily increasing demand for evidence of clinical effectiveness. It is believed that the limited amount of high-quality evidence is partly responsible for multiple geographic variations, inappropriate care, and also the limited success of quality improvement efforts (4, 5). As a result, the past decade has been marked by an unprecedented interest in evidence-based medicine and various types of avenues, including clinical practice guidelines, that can provide valid and reliable information about healthcare. Thus, healthcare decisions are purportedly and increasingly being made on research-based evidence rather than on expert opinion or clinical experience alone.

Systematic reviews and meta-analyses represent a rigorous method of compiling scientific evidence to answer questions regarding healthcare issues of treatment, diagnosis, or preventive services. Similar to meta-analysis and systematic reviews, another commonly used technique in the evaluation of the evidence is called health technology assessment (HTA). Practice guidelines provide another avenue in providing valid and reliable information about healthcare. They are systematically developed recommendations that assist the practitioner and the patient in making decisions about healthcare. They may be adapted, modified, or rejected based on healthcare needs or constraints.

Evidence-Based Medicine

Evidence-based medicine is defined as the conscientious, explicit and judicious use of the current best evidence in making decisions about the care of individual patients (6). Thus, evidence-based medicine is essentially what most clinicians have been trying to practice all their working lives. The practice of evidence-based medicine requires the integration of individual clinical expertise with the best available external evidence from systematic research. Decisions that affect the care of patients should be made with due weight accorded to *all* valid, relevant information. These include valid and relevant clinical evidence derived from randomized controlled trials, and *all* types of evidence, patient preferences, and resources. Thus, no one sort of evidence should necessarily be the determining factor in decision-making. *All* implies that there should be an active search for all that is valid, relevant information and that an assessment should be made of the accuracy of information and the applicability of the evidence to the decision in question (6). Four basic contingencies originally defined evidence-based practice (7).

- First, recognition of the patient's problem and construction of a structured clinical question.
- Second, the ability to efficiently and effectively search the medical literature to retrieve the best available

From Pain Management Center Of Paducah, Paducah, Kentucky, Massachusetts General Hospital Pain Center, and Department of Anesthesiology and Critical Care, Harvard Medical School, Boston, Massachusetts, and Department of Anesthesiology, University of Louisville School Of Medicine, Louisville, Kentucky
Address Correspondence: Laxmaiah Manchikanti, MD, 2831 Lone Oak Road, Paducah, Kentucky 42003
Disclaimer: There was no external funding in preparation of this manuscript.
Conflict of Interest: None
Acknowledgement:
Manuscript received on 12/1/04.
Revision submitted on 12/20/04
Accepted for publication on 12/29/04

evidence to answer the clinical question.
- Third, critical appraisal of the evidence.
- Fourth, integration of the evidence with all aspects of individual patient decision making to determine the best clinical care of the patient.

Thus, evidence-based medicine is a loose term which has been used based not only on the necessity to present a particular view, but also based on personal philosophy, bias and conjecture. This has led to a multitude of questions as to whether evidence-based medicine is truly based on evidence.

Clinical Practice Guidelines

In the 1990s, numerous guidelines were published in various countries around the world. Many professional organizations produced evidence-based or consensus guidelines (8-15), and the Cochrane collaboration of systematic reviews (16), which started in 1993, now has more than 3,000 collaborations worldwide. In pain management, the first so-called evidence-based guidelines were produced by the Agency for Health Care Policy Research (AHCPR) in 1994 (17). AHCPR produced 15 guidelines at a cost of $750,000,000, each at varying costs (18). The agency was eventually replaced with a small portion of its original budget and without the mandate to develop practice guidelines. AHCPR was renamed as The Agency for Healthcare Research and Quality (AHRQ). However, guideline development continued, experiencing an explosive growth with numerous publications appearing in the form of consensus statements, clinical guidelines, and books (8-21).

Clinical practice guidelines are systematically developed statements that assist clinicians, consumers and policy makers to make appropriate healthcare decisions. Such guidelines present statements of best "practice" based on a thorough evaluation of the evidence from published studies on the outcomes of treatment. The methods used for collecting and evaluating evidence and developing guidelines can be applied to a wide range of clinical interventions and disciplines, including interventional procedures, both diagnostic and therapeutic, pharmaceuticals, and others.

Patients and physicians increasingly seek to combine their personal beliefs about healthcare choices with attention to high-quality evidence in making individual decisions about care. Multiple medical professional societies produce guidelines to assist physicians and patients in making medical decisions (22). In the modern era, not only managed care organizations, but almost all healthcare insurers including federal programs are increasingly depending on systematic reviews and technology assessments to support quality improvement efforts and to develop coverage and payment policies (1, 23, 24). However, it is believed that the current clinical research in interventional pain management is not consistently producing an adequate supply of information to meet the needs of clinical and health policy decision makers. Tunis et al (1) described that the inability to address many common, important clinical questions in modern medicine suggests a systemic problem and creates knowledge gaps.

The methods of guideline development are complex. The development should ensure that treating patients according to the guidelines will achieve the outcomes that are desired. Guidelines can be developed for a wide range of subjects including interventional pain management and interventional techniques. Thus, interventional pain management guidelines are concerned with spinal pain, chronic pain or controlled substance usage and procedures including diagnostic and therapeutic interventional techniques. Given the multiple potential areas of interventional pain management, each area should be selected for guideline development. Guideline development depends on the integration of a number of activities, from collection and processing of scientific literature to evaluation of the evidence, development of evidence-based recommendations, and implementation and dissemination of the guidelines to relevant professionals and consumers.

Guidelines are designed to improve the quality of healthcare and decrease the use of unnecessary, ineffective or harmful interventions. In an era of evidence-based medicine, guidelines are becoming one of the critical links between the best available evidence and good clinical practice. Guidelines constitute one element of a systematic approach to quality healthcare. Research has shown that clinical practice guidelines can be an effective means of changing the process of healthcare and improving health outcomes (25-28). Traditionally, guidelines have been based on consensus among experts. However, now it has been acknowledged that guideline recommendations should be based on systematic identification and synthesis of the best available scientific evidence (25, 29). Considering the extensive research activity, the lack of a single source to identify the appropriate literature, significant bias in the systematic evaluations, and substantial reports outside the published and peer reviewed literature; identification and synthesis of the available evidence; and publication of this evidence in the form of guidelines can be a major undertaking. The National Health and Medical Research Council of Australia published: A Guide to the Development, Implementation and the Evaluation of Clinical Practice Guidelines in 1999 (29). This comprehensive document includes 9 basic principles, the development, the dissemination and implementation of guidelines as shown in Table 1.

Shekelle et al (30) presented five initial steps in the development of an evidence-based guideline.

1. Identification and refinement of the subject area
2. Convening and running guideline development groups
3. Assessment of the evidence about the clinical question or condition
4. Translation into a recommendation within a clinical practice guideline
5. External review of the guideline

They also described other steps which included the dissemination, implementation, and evaluation with practice guidelines.

Legal considerations and potential liability of practitioners is an important aspect of guidelines. Many practitioners are concerned about their potential legal liability if a patient does not receive treatment as specified in clinical practice guidelines. It is possible that guidelines could be produced as evidence of what constitutes reasonable conduct by an interventional pain management practitioner. It is generally believed that following the guidelines, provides a measure of protection. However, physicians should provide all appropriate information about all types of treatments, along with associated risks of any treatment, especially risks that may influence the patient's decision. Patients should be provided with as much information as they seek, and in a form that is appropriate to their culture and level of

Table 1. *Basic principles described for development of guidelines*

1.	Focus on outcomes ♦Survival rates to quality-of-life attributes
2.	Best available evidence ♦ Graded according to its quality, relevance and strength
3.	Appropriate systems to synthesize the available evidence ♦ Turning the evidence into a clinically useful recommendation depends on the judgment, experience and good sense of the authors of guidelines. ♦ **The fact of having evidence from a high level study does not automatically result in a good clinical recommendation**.
4.	Multidisciplinary process of development.
5.	Flexibility and adaptability
6.	Evaluation of cost effectiveness of treatments
7.	Appropriate dissemination
8.	Evaluation of implementation and impact of guidelines
9.	Appropriate revision of the guidelines on a regular basis

Adapted from Reference 29

education. Finally, all patients should be encouraged to make their own decisions. The potential for any guidelines to be used as evidence in a court of law depends on the process used to develop them, the extent to which they are evidence-based, the degree of consensus about them, and whether they are up to date (29). However, guideline developers are unlikely to be held liable for any negative consequences of the implementation of guidelines. In general, guidelines should be summaries of the evidence, should have an expiration date, should not be unduly prescriptive, and should acknowledge areas where there is disagreement.

Shaneyfelt et al (31) reviewed the methodological quality of clinical guidelines in the peer-reviewed medical literature, with evaluation of 279 guidelines developed by 69 different organizations, published from 1985 to 1997. They showed that mean overall adherence to standards by each guideline was 43.1%. They concluded that guidelines in the peer-reviewed medical literature during the past decade did not adhere well to published methodological standards. They also added that while all areas of guideline development need improvement, the greatest improvement is needed in the identification, evaluation, and synthesis of the scientific evidence. The standards described are listed in Table 2.

Reviews in Interventional Pain Management

Most systematic evaluations in interventional pain management included only randomized controlled trials. This is in contrast to the definition of evidence-based medicine, which explicitly states that no one sort of evidence should necessarily be the determining factor in decision-making. Further, evidence-based medicine also emphasizes *all*, implying that there should be an active search for *all* that is valid, relevant information and that assessment should be made of the accuracy of information and the applicability of the evidence to the decision in question. Recent systematic analyses have increasingly utilized observational studies, as well as other types of evidence, even though this approach has not been applied to interventional pain management (32). It is also recognized that meta-analysis restricted to randomized clinical trials is usually preferred to meta-analysis of observational studies (33-35). However, this has not been demonstrated in interventional pain management. In many situations, randomized designs are not feasible, and only data from observational studies are available (36).

Practical Clinical Trials

Tunis et al (1) reported that the prevalence and significance of gaps in knowledge about clinical effectiveness are most readily appreciated by reviewing the results of most systematic literature reviews, technology assessments, and clinical practice guidelines. A consistent finding of most reviews appears to be that the quality of evidence available to answer the critical questions identified by experts is suboptimal. As an example, most systematic reviews performed in interventional pain management include studies providing data not applicable to patients treated in typical practice settings. Consequently, organizations that develop evidence-based clinical practice guidelines may not be able to develop clear, specific recommendations. The limited quantity and quality of available scientific information impedes the efforts of public and private health insurers in developing evidence-based coverage policies for many new and existing technologies (24, 37). Further, poor quality studies of new technologies can lead to inappropriate spending being allocated for new technologies for which the long-term benefits and risks have not been determined, with an unintended effect of reducing the more effective old technologies.

Trials of healthcare interventions are often described as either explanatory or pragmatic (38). Explanatory trials generally measure efficacy – the benefit a treatment produces under ideal conditions. Consequently, explanatory trials often use carefully defined subjects in a well-controlled research setting. In contrast, pragmatic trials, also known as practical clinical trials, measure effectiveness – the benefit the treatment produces in routine clinical practice.

Patient selection in an explanatory approach is based on the principles of homogenous population, primarily aiming to further scientific knowledge. However, in a pragmatic or practical clinical trial, the design reflects variations between patients that occur in real life clinical settings, and aims to inform choices between treatments. Consequently, to ensure generalizability, pragmatic trials should, so far as possible, represent the patients to whom the treatment will be applied (38).

The next aspect relates to randomization, which is the focus of clinical research which deals with selection bias. Even then, multiple other sources of bias may affect the results. However, biased assessment of outcome has to be dealt with both in explanatory, as well as pragmatic trials by having an independent assessor who is blind to treatment allocation. Biases resulting from a pragmatic or practical clinical trial are accepted as part of physician's and patient's responses to treatment and are considered in the overall assessment. Consequently, in pragmatic approaches, the treatment response is the total difference between two treatments, including both treatment and associated placebo effects, as this will best reflect the likely clin-

Table 2. *Standards for guidelines*

- Purpose of the guidelines is specified
- Rationale and importance of the guideline are explained
- The participants in the guideline development process and their areas of expertise are specified
- Targeted health problem or technology is clearly defined
- Targeted patient population is specified
- Intended audience or users of the guideline are specified
- The principal preventive, diagnostic, or therapeutic options available to clinicians and patients are specified
- The health outcomes are specified
- The method by which the guideline underwent external review is specified
- An expiration date or date of scheduled review is specified
- Method of identifying scientific evidence is specified
- Time period from which evidence is reviewed is specified
- The evidence used is identified by citation and referenced
- Method of data extraction is specified
- Method for grading or classifying the scientific evidence is specified
- Formal methods of combining evidence or expert opinion are used and described
- Benefits and harms of specific health practices are specified
- Benefits and harms are quantified
- The effect on health care costs from specific health practices is specified
- Costs are quantified
- The role of value judgments used by the guideline developers in making recommendations is discussed
- The role of patient preferences is discussed
- Recommendations are specific and apply to the stated goals of the guideline
- Recommendations are graded according to the strength of the evidence
- Flexibility in the recommendations is specified

Adapted from Reference 31

ical response in practice (39-46). Thus, placebo-controlled trials take a different approach (47-52).

The practical clinical trials best address questions about the risks, benefits, and costs of an intervention as they would occur in routine clinical practice (38). Thus, it is thought that the most distinctive features of practical clinical trials are that they select clinically relevant interventions to compare, include a diverse population of study participants, recruit participants from a variety of practice settings, and collect data on a broad range of health outcomes, simulating actual practices.

Practical clinical trials often are designed to compare viable alternative clinical strategies. A practical clinical trial of acute low back pain randomized 323 patients to one of three widely used treatments which included physical therapy, chiropractic treatment, and self-care from an educational booklet, and showed that physical therapy and chiropractic care increased patient satisfaction and marginally reduced symptoms compared with the self-care principles outlined in the booklet (53). A practical clinical trial of therapeutic massage compared with acupuncture also demonstrated that therapeutic massage was more effective and less costly than acupuncture in treating low back pain (54). In interventional pain management, multiple practical clinical trials have been performed evaluating the effectiveness of less expensive and safer modalities compared to more invasive treatments (39-46). However, these studies have been appreciated neither by insurers nor by academicians focusing on placebo-controlled treatments.

Randomized Trials

The acme of clinical research is the randomized, double blind, controlled trials, but such trials must be undertaken responsibly and are extremely difficult to conduct in interventional pain management. Randomized controlled trials were introduced into clinical medicine when streptomycin was evaluated in the treatment of tuberculosis (55). Since then, randomized controlled trials have become the gold standard for assessing the effectiveness of therapeutic agents (56-58). Comparison of published randomized controlled studies with those that used observational designs concluded that bias in patient selection may irretrievably weigh the outcome of historically controlled trials in favor of new therapies in observational studies (59). However, others (60, 61) have concluded that randomized clinical trials, significantly favored experimental interventions. Many stumbling blocks, including the issues of ethics, feasibility, cost and reliability, and insurmountable challenges to randomized, double-blind trials in interventional pain management have been discussed (62-67).

The presumed unsurpassed science of a randomized trial is attributed to the ability to assign subjects randomly. However, random assignment does not confer an absolute protection against bias. It simply reduces the likelihood that such bias has occurred. Because randomized controlled trials are complicated and difficult to conduct, they are usually restricted to very tightly targeted groups of patients. Often, the investigators are not actively concerned about how subjects are obtained and rely on random allocation to distribute any differences equally across the two groups. As a result, randomized trials often trade internal validity (tightness of comparisons) for external validity (generalizability) (68). Generally, a difference in outcome between a treatment and a control group can be due to chance, confounding, or bias due to differences between the groups, differences in handling the groups, and the true effect of intervention. Confounding and bias are avoided in the design of a trial by randomization, single-blinding or double-blinding. Assigning volunteers to the treatment group and those who do not volunteer to the control group is also likely to result in a biased comparison - volunteers will be quite different, in many respects, from patients who do not volunteer (69). The criticism also has been advanced against allocation and treatment based on alternate days, alternate numbers or another assigned preformed methodology. In evalu-

Table 3. *Differences between narrative and systematic reviews*

Core Feature	Narrative Reviews	Systematic Reviews
Study Question	Often broad in scope.	Often a focused clinical question.
Data sources and search strategy	Specifications of database searched and search strategy are not typically provided.	Comprehensive search of many databases as well as the so-called "gray literature". Explicit search strategy provided.
Selection of articles for study	Not usually specified. If specified, potentially biased.	Criterion-based selection, uniformly applied.
Article review or appraisal	Variable, depending on who is conducting the review.	Rigorous critical appraisal, typically using a data extraction form.
Study quality	Usually not assessed. If assessed, may not use formal quality assessment.	Some assessment of quality is almost always included as part of the data extraction process.
Synthesis	Often a qualitative summary.	Quantitative or qualitative summary.
Inferences	Occasionally evidence-based.	Usually evidence-based.

Adapted from Reference 73

ating the influence of randomization over other types of allocation research designs in interventional pain management, it has been shown that allocation without randomization was homogenous and superior to randomization.

Non-Randomized or Observational Studies

The role of observational studies in the evaluation of treatment is a long-standing and contentious topic (16, 20, 32). However, any systematic review of evidence on a therapeutic topic needs to take into account the quality of the evidence. Observational studies similar to randomized trials may have flaws in design or analysis (16, 20, 32). The problems of heterogeneity and publication bias are relevant to all comparisons of evidence from randomized, controlled studies, as well as observational studies. However, all observational studies have one crucial deficiency: lack of experimental design. Multiple reviews have been performed comparing the evidence of treatment effects in randomized and non-randomized studies (70-72). These evaluations have suggested that for selected medical topics, both randomized and non-randomized studies, may yield very similar results.

Evaluation of the Trials

Throughout the 1990s and into the 21st century, the Agency for Health Care Research and Quality (AHRQ) has been the foremost federal agency providing research support and policy guidance in health services research in the United States (73). This comprehensive document includes methodology and results of systems for rating the quality of individual articles, as well as systems for grading the strength of a body of evidence.

The National Health and Medical Research Council of Australia considered scientific evidence to be at the core of evidence-based approach to clinical or public health issues (74). They emphasized that evidence needs to be carefully gathered and collated from a systematic literature review of a particular issue in question.

The National Coordinating Center for Health Technology Assessment of the United Kingdom also published systematic reviews of trials and other studies (75). They described extensively the methodology for meta-analysis, searching the literature and identifying primary studies, evaluating the study quality, applications of meta-analysis in other contexts and using other data types, extensions of meta-analytic methods, and recommendations for further research.

Cochrane collaboration (16) has also advanced many principles of evidence synthesis. In fact, Cochrane reviews include multiple interventional techniques including injection therapy in managing low back pain.

Types of Reviews

A systematic review is a type of scientific investigation of the literature on a given topic in which the "subjects" are the articles being evaluated (76). In contrast, a narrative review is similar to a systematic review but without all the safeguards to control against bias (Table 3). The major differences between these two approaches to synthesizing the clinical or scientif-

Table 4. *Differences between systematic reviews and health technology assessments (HTAs)*

	Systematic Reviews	Health Technology Assessments
Methodological standards	Only include studies with the best methodological evidence	Include studies of topics of interest to policy-makers, even if evidence is suboptimal
Repeating previous studies	No need to repeat if previous studies were high quality, and no new high-quality evidence	The need to defend the report's conclusions often necessitates repetition
Breadth versus depth	Only include topics for which there is good evidence; topics driven by scientists' interests	Include topics most relevant to policy-makers; exclude those not of relevance even if there is good quality evidence
Inclusion of content experts and policy-makers	Content experts, but not policy-makers usually included	Can be concerns that content experts and policy-makers are biased
Performance of economic evaluations	Usually not done	Economic evaluations are an important component of HTAs, but lack of good evidence about effectiveness/diagnostic accuracy limit their impact
Making policy recommendations	Almost never done	Sometimes done, but with caution
Active dissemination	Rarely done	Sometimes done

Adapted from Reference 83

ic literature is that a systematic review attempts to minimize bias by the comprehensiveness and reproducibility of the search for and selection of articles for review, and provides methodologic quality studies (77-82).

A third type of review is Health Technology Assessment (HTA), a multidisciplinary field that studies the medical, social, ethical, and economic implications of the development, use, and diffusion of health technologies (83). HTA has been described as, "the bridge between the world of research and the world of decision making." HTAs are being performed with increasing frequency and have influenced decision-making in many jurisdictions. To effectively influence policy-makers, the authors of HTAs must not only strive for scientific accuracy, but must also be aware of other issues such as the optimal timing of the reports released, their political sensitivity to the important decision-makers, and how best to disseminate the results. Differences between systematic reviews and health technology assessments are illustrated in Table 4.

The Quality Of Individual Articles

Multiple types of studies used for assessing clinical and public health interventions are described in Table 5, which include systematic reviews, experimental studies, randomized trials, observational studies, and diagnostic test studies (74).

AHRQ described important evaluation domains and elements for evaluating systems related to rating the quality of individual articles, including systematic reviews, randomized clinical trials, observational studies and diagnostic test studies (73). Cochrane Review Group (84) also described methodology for systemic reviews for spinal disorders. Additional guidelines have been developed to evaluate accuracy of diagnostic studies (85). Further, various types of reporting guidelines such as the CONSORT (86), QUOROM (87), STARD (88), and MOOSE (32) have been described in reporting various individual articles. These guidelines assist the reporting authors by reducing the unavoidable tension and by facilitating easier publication . Tables 6 to 11 show the important domains and elements for systems to rate quality of individual articles.

Systematic Reviews

Authors of this AHRQ document reviewed 20 systems concerned with systematic reviews or meta-analyses. To arrive at a set of high-performing scales or checklists pertaining to systematic reviews, the authors of AHRQ (73) took account of 7 key domains as shown in Table 6: study question, search strategy, inclusion and exclusion criteria, data extraction, and funding or sponsorship.

Randomized Clinical Trials

For evaluation of randomized trials, two types of guidelines are available. These include the guidelines described by AHRQ (73) and other commonly used guidelines in evaluation of interventions described by Cochrane Review Group (84).

The authors of AHRQ designated a set of high-performing scales or checklists pertaining to randomized clinical trials

Table 5. *Types of studies used for assessing clinical and public health interventions*

Study design	Protocol
Systematic reviews	Systematic location, appraisal and synthesis of evidence from scientific studies
Experimental studies	
Randomized controlled trial	Subjects are randomly allocated to groups for either the intervention/treatment being studied or control/placebo (using a random mechanism, such as coin toss, random number table, or computer-generated random numbers) and the outcomes are compared.
Pseudorandomized controlled trial	Subjects are allocated to groups for intervention/treatment or control/placebo using a nonrandom method (such as alternate allocation, allocation by days of the week, or odd-even study numbers) and the outcomes are compared.
Clustered randomized trial	Groups of subjects are randomized to intervention or control groups (eg, community intervention trials).
Comparative (nonrandomized and observational) studies	
Concurrent control or cohort	Outcomes are compared for a group receiving the treatment/intervention being studied, concurrently with control subjects receiving the comparison treatment/intervention (eg, usual or no care).
Case-control	Subjects with the outcome or disease and an appropriate group of controls without the outcome or disease are selected and information is obtained about the previous exposure to the treatment/intervention or other factor being studied.
Historical control	Outcomes for a prospectively collected group of subjects exposed to the new treatment/intervention are compared with either a previously published series or previously treated subjects at the same institutions.
Interrupted time series	Trends in the outcome or disease are compared over multiple time points before and after the introduction of the treatment/intervention or other factor being studied.
Other observational studies	
Case series	A single group of subjects are exposed to the treatment/intervention.
-- post-test	Only outcomes after the intervention are recorded in the case series, so no comparisons can be made.
-- pretest/post-test	Outcomes are measured in subjects before and after exposure to the treatment/intervention for comparison (also called a 'before-and-after' study).

Adapted from Reference 74

Table 6. *Domains and elements for systematic reviews*

Domain[#]	Elements*
Study Question	• **Question clearly specified and appropriate**
Search Strategy	• ***Sufficiently comprehensive and rigorous with attention to possible publication biases***
	• *Search restrictions justified (e.g., language or country of origin)*
	• Documentation of search terms and databases used
	• Sufficiently detailed to reproduce study
Inclusion and Exclusion Criteria	• **Selection methods specified and appropriate, with a priori criteria specified if possible**
Interventions	• **Intervention(s) clearly detailed for all study groups**
Outcomes	• **All potentially important harms and benefits considered**
Data Extraction†	• Rigor and consistency of process
	• Number and types of reviews
	• Blinding of reviewers
	• Measure of agreement or reproducibility
	• Extraction of clearly defined interventions/exposures and outcomes for all relevant subjects and subgroups
Study Quality and Validity	• ***Assessment method specified and appropriate***
	• Method of incorporation specified and appropriate
Data Synthesis and Analysis	• ***Appropriate use of qualitative and/or quantitative synthesis, with consideration of the robustness of results and heterogeneity issues***
	• Presentation of key primary study elements sufficient for critical appraisal and replication
Results	• **Narrative summary and/or quantitative summary statistic and measure of precision, as appropriate**
Discussion	• **Conclusions supported by results with possible biases and limitations taken into consideration**
Funding or Sponsorship	• ***Type and sources of support for study***

[#] Key domains are in italics; *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.; Adapted from ref 73

Table 7. *Domains and elements for randomized controlled trials*

Domain[#]	Elements*
Study Question	• **Clearly focused and appropriate question**
Study Population	• **Description of study population**
	• **Specific inclusion and exclusion criteria**
	• Sample size justification
Randomization	• *Adequate approach to sequence generation*
	• ***Adequate concealment method used***
	• *Similarity of groups at baseline*
Blinding	• ***Double-blinding (e.g., of investigators, caregivers, subjects, assessors, and other key study personnel as appropriate) to treatment allocation***
Interventions	• **Intervention(s) clearly detailed for all study groups (e.g., dose, route, timing for drugs, and details sufficient for assessment and reproducibility for other types of interventions)**
	• Compliance with intervention
	• Equal treatment of groups except for intervention
Outcomes	• **Primary and secondary outcome measures specified**
	• Assessment method standard, valid, and reliable
Statistical Analysis	• ***Appropriate analytic techniques that address study withdrawals, loss to follow-up, missing data, and intention to treat***
	• Power calculation
	• Assessment of confounding factors
	• Assessment of heterogeneity, if applicable
Results	• **Measure of effect for outcomes and appropriate measure of precision**
	• Proportion of eligible subjects recruited into study and followed up at each assessment
Discussion	• **Conclusions supported by results with possible biases and limitations taken into consideration**
Funding or Sponsorship	• ***Type and sources of support for study***

[#] Key domains are in italics; *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. Adapted from ref 73

by assessing their coverage of the following 7 domains as shown in Table 7; study population, randomization, blinding, interventions, outcomes, statistical analysis and funding or sponsorship. The criteria described by the Cochrane Review Group for musculoskeletal disorders are illustrated in Table 8.

Observational Studies

Cochrane collaboration and AHRQ recognize the importance of observational studies. Authors of AHRQ considered several key domains and arrived at a set of 5 high-performing scales or checklists pertaining to observational studies as described in Table 9: comparability of subjects, exposure or intervention, outcome measurement, statistical analysis and funding or sponsorship. Apparently the systems that cover these domains represent acceptable approaches for assessing the quality of observational studies.

Cochrane Reviewers Handbook 4.2.0 described multiple advantages and dangers of including non-randomized studies in systematic reviews, along with guidelines on how to do these.

The advantages include that if a systematic review relies solely on data from a randomized trial, it is open to a number of problems. The most obvious of these is that certain important health care problems have not been studied, or are impossible or very difficult to study in randomized trials. Randomized trials may be inadequate for other reasons also. For example, there may be insufficient information on the types of participants or outcome which are of relevance to the review (e.g., rare side effects), or the data may only contain short-term follow-up when important findings depends on longer follow-up. Inclusion of evidence from non-randomized studies may resolve some of these problems.

However, inclusion of non-randomized studies in systematic reviews may also pose problems and a threat to validity as unexpected biases may creep in and invalidate the conclusions.

Studies of Diagnostic Tests

Multiple precision diagnostic blocks utilized in interventional pain management have never been reviewed systematically except for facet joint blocks (89). However, the value and validity of multiple diagnostic interventions with precision diagnostic blocks has been described

Table 8. *Methodologic quality criteria of internal validity list of Cochrane Musculoskeletal Review Group*

Patient selection			
1. Treatment allocation			
Was the method of randomization described and adequate?	Yes	No	Don't know
Was the treatment allocation concealed?	Yes	No	Don't know
2. *Were the groups similar at baseline regarding the most important prognostic indicators?*	Yes	No	Don't know
Intervention			
3. Was the care provider blinded?	Yes	No	Don't know
4. Was controlled for co-interventions which could explain the results?	Yes	No	Don't know
5. Was the compliance rate (in each group) unlikely to cause bias?	Yes	No	Don't know
6. Was the patient blinded?	Yes	No	Don't know
Outcome measurement			
7. *Was the outcome assessor blinded?*	Yes	No	Don't know
8. *Was at least one of the primary outcome measures applied?*	Yes	No	Don't know
9. *Was the withdrawal/drop-out rate unlikely to cause bias?*	Yes	No	Don't know
Statistics			
10. Did the analysis include an intention-to-treat analysis?	Yes	No	Don't know

Adapted from Reference 84

extensively and also has been questioned repeatedly.

AHRQ Assessment identified 6 checklists to evaluate the quality of diagnostic studies. The authors identified 5 key domains for making judgments about the quality of diagnostic test reports as described in Table 10: study population, adequate description of the test, appropriate reference standard, blinded comparison of test and reference and avoidance of verification bias.

Due to multiple difficulties with assessment of quality of diagnostic studies, a new tool known as QUADAS, has been described (85). These items are illustrated in Table 11. This instrument fills a gap in systemic evaluation of diagnostic accuracy studies.

Grading the Strength of Body of Evidence

Systems for grading the strength of a body of evidence are much less uniform and consistent than those for rating study quality (73). As with the quality rating systems, selecting among the evidence grading systems will depend on the reason for measuring evidence strength, the type of studies that are being summarized, and the structure of the review panel. Domains for rating the overall strength of a body of evidence are listed in Table 12. The National Health and Medical Research Council (NHMRC) (74) described five key points for considering levels of evidence as follows as listed in Table 13. Some systems are extremely cumbersome to use, requiring substantial resources, whereas others are incomplete and incomprehensive. Multiple systems have been utilized in preparation of guidelines. Table 14 shows the designation of levels of evidence from level I through V considered in interventional pain management with guideline preparation (9, 21).

Searching for the Evidence

To achieve balance in evidence-based interventional pain management and also to include all types of evidence, *all* types of evidence must be literally included. These include not only systematic reviews and randomized clinical trials but also all published literature of observational studies and diagnostic test studies. Thus, a search strategy should include all sources easily available to obtain the literature.

It has been shown that using only MEDLINE, 30% to 80% of all known published randomized controlled trials are identifiable, depending on the area or specific question. In systematic reviews of trials and other studies, it was described that non-English language references are underrepresented in MEDLINE and only published articles are included. Thus, there is the potential for publication bias and language bias. Further, it was shown that depending on the country of origin, there is also potential for geographical biases. Another problem with databases is that even though many of the studies may be included in a database such as MEDLINE, it may not be easy to identify all those which are relevant. MEDLINE failed to find 44% of known trials. Possible reasons for poor retrieval are as follows: the search used was too narrow, the indexing of studies in MEDLINE is inadequate and the original reports may have been too vague. The same issues are applicable to EMBASE. In general, MEDLINE provides wide coverage of many English language journals. In contrast, EMBASE can be used to increase coverage of articles in the European languages. The overlap between the MEDLINE and EMBASE is approximately 34%, even though it can vary between 10% and 75% for specific topics. Thus, one cannot rely on searching a single database. Further, dependence on databases also may miss many non-indexed journals, proceedings of the scientific meetings, and peer-reviewed articles from scientific newsletters. Search of the reference lists of articles found through databases may also identify further studies for consideration. In fact, the Cochrane

Table 9. *Domains and elements for observational studies*

Domain[#]	Elements
Study Question	• **Clearly focused and appropriate question**
Study Population	• **Description of study populations** • Sample size justification
Comparability of Subjects†	For all observational studies: • Specific inclusion/exclusion criteria for all groups • Criteria applied equally to all groups • Comparability of groups at baseline with regard to disease status and prognostic factors • Study groups comparable to non-participants with regard to confounding factors • *Use of concurrent controls* • Comparability of follow-up among groups at each assessment Additional criteria for case-control studies: • Explicit case definition • Case ascertainment not influenced by exposure status • Controls similar to cases except without condition of interest and with equal opportunity for exposure
Exposure or Intervention	• **Clear definition of exposure** • Measurement method standard, valid and reliable • Exposure measured equally in all study groups
Outcome Measurement	• **Primary/secondary outcomes clearly defined** • Outcomes assessed blind to exposure or intervention status • Method of outcome assessment standard, valid and reliable • Length of follow-up adequate for question
Statistical Analysis	• Statistical tests appropriate • Multiple comparisons taken into consideration • Modeling and multivariate techniques appropriate • Power calculation provided • **Assessment of confounding factors** • Dose-response assessment, if appropriate
Results	• **Measure of effect for outcomes and appropriate measure of precision** • Adequacy of follow-up for each study group
Discussion	• **Conclusions supported by results with possible biases and limitations taken into consideration**
Funding or Sponsorship	• ***Type and sources of support for study***

[#] Key domains are in italics
*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.
Adapted from ref 73

Table 10. *Domains and elements for diagnostic studies*

Domain[#]	Elements*
Study Population	• ***Subjects similar to populations in which the test would be used and with a similar spectrum of disease***
Adequate Description of Test	• ***Details of test and its administration sufficient to allow for replication of study***
Appropriate Reference Standard	• ***Appropriate reference standard ("gold standard") used for comparison*** • Reference standard reproducible
Blinded Comparison of Test and Reference	• Evaluation of test without knowledge of disease status, if possible • ***Independent, blind interpretation of test and reference***
Avoidance of Verification Bias	• ***Decision to perform reference standard not dependent on results of test under study***

[#] Key domains are in italics
*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.
Adapted from ref 73

Table 11. *Items utilized for assessment of quality of individual articles of diagnostic studies by QUADAS tool*

Item		Yes	No	Unclear
1.	Was the spectrum of patients representative of the patients who will receive the test in practice?	( )	( )	( )
2.	Were selection criteria clearly described?	( )	( )	( )
3.	Is the reference standard likely to correctly classify the target condition?	( )	( )	( )
4.	Is the time period between reference standard and index test short enough to be reasonably sure that the target condition did not change between the two tests?	( )	( )	( )
5.	Did the whole sample or a random selection of the sample, receive verification using a reference standard of diagnosis?	( )	( )	( )
6.	Did patients receive the same reference standard regardless of the index test result?	( )	( )	( )
7.	Was the reference standard independent of the index test(i.e. the index test did not form part of the reference standard)?	( )	( )	( )
8.	Was the execution of the index test described in sufficient detail to permit replication of the test?	( )	( )	( )
9.	Was the execution of the reference standard described in sufficient detail to permit its replication?	( )	( )	( )
10.	Were the index test results interpreted without knowledge of the results of the reference standard?	( )	( )	( )
11.	Were the reference standard results interpreted without knowledge of the results of the index test?	( )	( )	( )
12.	Were the same clinical data available when test results were interpreted as would be available when the test is used in practice?	( )	( )	( )
13.	Were uninterpretable/intermediate test results reported?	( )	( )	( )
14.	Were withdrawals from the study explained?	( )	( )	( )

Adapted from Reference 85

Table 12. *Domains for rating the overall strength of a body of evidence*

Domain	Definition
Quality	• The quality of all relevant studies for a given topic, where "quality" is defined as the extent to which a study's design, conduct and analysis has minimized selection, measurement, and confounding biases
Quantity	• The magnitude of treatment effect
	• The number of studies that have evaluated the given topic
	• The overall sample size across all included studies
Consistency	• For any given topic, the extent to which similar findings are reported from work using similar and different study designs.

Adapted from Reference 73

Table 13. *Keypoints in consideration of level of evidence*

- ♦ Resolution of differences in the conclusions reached about effectiveness from studies at differing levels of evidence or within a given level of evidence.
- ♦ Resolution of the discrepancies is an important task in the compilation of an evidence summary.
- ♦ Inclusion of biostatistical and epidemiological advice on how to search for possible explanation for the disagreements before data are rejected as being an unsuitable basis on which to make recommendations.
- ♦ Recognition of the fact that it may not be feasible to undertake randomized controlled trials in all situations. Guidelines should be used on the best available evidence.
- ♦ Recognition of the fact that it may be necessary to use evidence from different study designs for different aspects of the treatment effect.

Adapted from Reference 74

Table 14. *Designation of levels of evidence*

Level I	Conclusive: Research-based evidence with multiple relevant and high-quality scientific studies or consistent reviews of meta-analyses
Level II	Strong: Research-based evidence from at least one properly designed randomized, controlled trial; or research-based evidence from multiple properly designed studies of smaller size; or multiple low quality trials.
Level III	Moderate: a) Evidence obtained from well-designed pseudo-randomized controlled trials (alternate allocation or some other method); b) evidence obtained from comparative studies with concurrent controls and allocation not randomized (cohort studies, case-controlled studies, or interrupted time series with a control group); c) evidence obtained from comparative studies with historical control, two or more single-arm studies, or interrupted time series without a parallel control group.
Level IV	Limited: Evidence from well-designed non-experimental studies from more than one center or research group; or conflicting evidence with inconsistent findings in multiple trials
Level V	Indeterminate: Opinions of respected authorities, based on clinical evidence, descriptive studies, or reports of expert committees.

Adapted and modified from Reference 9

handbook advises that reviewers should check the references of all relevant articles that are obtained. Thus, additional potentially relevant, articles that are identified should be retrieved and assessed for possible inclusion in the review. The potential for reference bias or a tendency to potentially cite studies supporting one's own views, however, should be kept in mind when doing this type of search. This bias can be guarded against by using a multitude of search strategies.

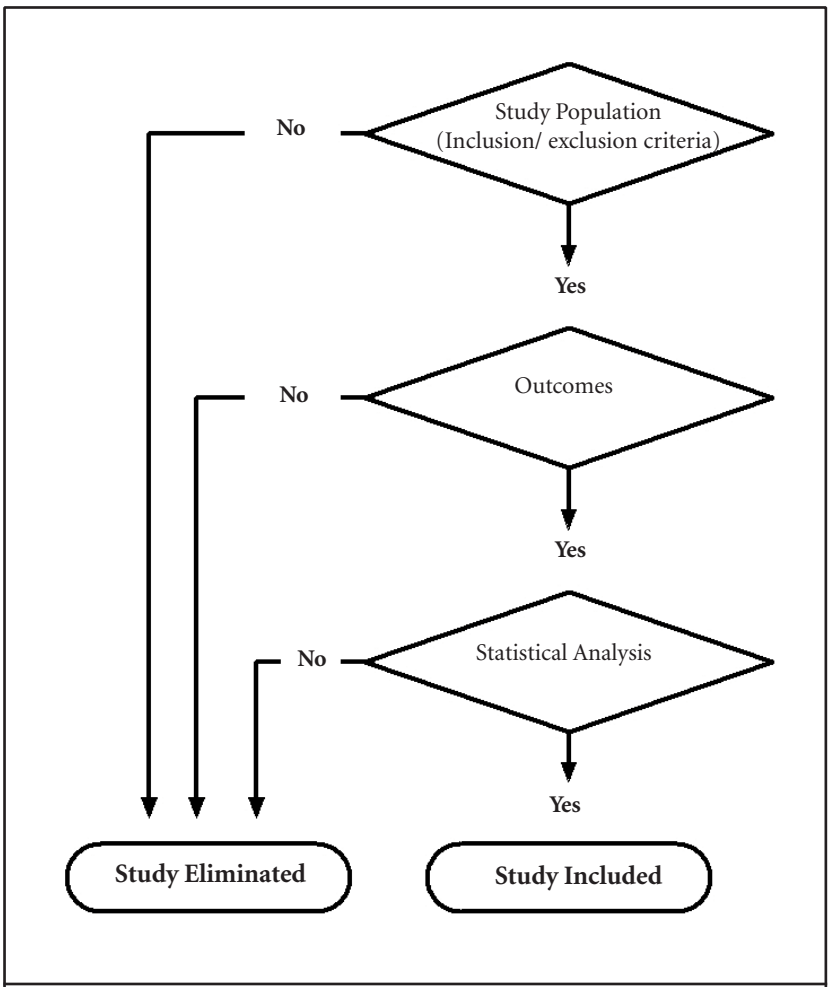


Fig. 1. *Study evaluation (inclusion/exclusion) algorithm*

Inclusion/Exclusion Criteria

Inclusion and exclusion criteria must be established for any type of review. Two types of inclusion/exclusion criteria are considered in review of articles for interventional pain management (9, 21, 88, 90, 91). Applicable inclusion/exclusion criteria and an algorithm are shown in Table 15 and Fig. 1.

To meet the inclusion criteria, a study should answer at least one question positively in all three categories.

Analytical Preparation

The evidence linkages or synthesis are performed by systematic reviews, as well as meta-analysis. In both types of analyses, methodological criteria and controls are crucial (Fig. 2). Apart from these two, consensus are also utilized as evidence. Research findings from published literature provide the cornerstone for guideline recommendations. However, published studies alone may not provide all the necessary or complete information regarding details of clinical practice of interventional techniques. Consequently, additional sources of information and evidence, as well as consensus are sought. The consensus data generally is obtained from the guideline committee through the members of the committee or it may be extended to other experts in the field or by open forum presentations.

Guideline recommendations generally are based directly on the evidence linkages developed during the process. Generally, all sources of evidence, including systematic reviews and consensus are utilized. However, they are separately considered and provided with different weights prior to formulation of the final recommendations.

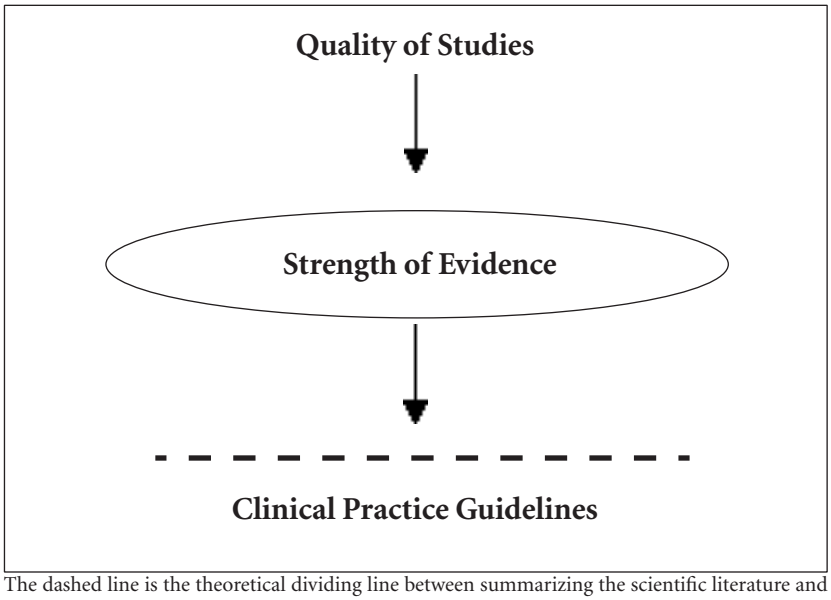
Table 15. *Inclusion and exclusion criteria*

1. Are the patients described in sufficient detail to allow you to decide whether they are comparable to those that are seen in clinical practices of interventional pain management?
   - A) Setting – office, hospital, outpatient, inpatient
   - B) Physician – interventional pain physician, general physician, anesthesiologist, physiatrist, neurologist, rheumatologist, orthopedic surgeon, neurosurgeon, etc.
   - C) Patient characteristics - duration of pain
   - D) Non-interventional techniques or surgical intervention in the past
   - E) Exclusion criteria
   - F) Inclusion criteria
2. Is the intervention described well enough to enable you to provide the same for patients in interventional pain management settings?
   - A) Nature of intervention
   - B) Frequency of intervention
   - C) Duration of intervention
3. Were clinically relevant outcomes measured?
   - A) Proportion of pain relief
   - B) Disorder/specific disability
   - C) Functional improvement
   - D) Allocation of eligible and non-eligible patients to return to work
   - E) Ability to work
   - F) Psychological assessment or improvement

Adapted and modified from Reference 90, 91

Conclusion

Evidence synthesis and the development of guidelines in interventional pain management is a complex and difficult task. Evidence synthesis is performed by evaluating all types of evidence based on the evidence-based principles. Practice guidelines are systematically developed to assist the practitioner and the patient in making decisions about healthcare. Practice recommendations may be adapted, modified, or rejected according to clinical needs and constraints. Consequently, practice guidelines are not intended as standards or absolute requirements. This review has described evidence-based medicine and the systematic process of evidence synthesis and development of guidelines in interventional pain management in practical terms.



The dashed line is the theoretical dividing line between summarizing the scientific literature and developing a clinical practice guideline. Below the dashed line, guideline developers would decide whether the evidence represents all the relevant subsets of the populations (or settings, or types of clinicians) for whom the guideline is being developed.
Adapted from Reference 73

Fig. 2. *Continuum from study quality through strength of evidence to guideline development*

Author Affiliation:

Laxmaiah Manchikanti, MD
Medical Director
Pain Management Center of Paducah
2831 Lone Oak Road
Paducah, Kentucky 42003
Assistant Clinical Professor of Anesthesiology and Perioperative Medicine
University of Louisville, Kentucky 40292
E-mail: drm@apex.net

Salahadin Abdi, MD, PhD
Director, Massachusetts General Hospital Pain Center
Department of Anesthesiology and Critical Care
15 Parkman street, Suite 333B
Boston, MA 02114
Department of Anesthesiology and Critical Care
Harvard Medical School
Boston, Massachusetts
E-mail: sabdi@partners.org

Linda F. Lucas, MD
Associate Professor
Department of Anesthesiology and Perioperative Medicine
Director of Pain Management Fellowship
University of Louisville School of Medicine
530 S. Jackson Street
Louisville, Kentucky 40202
E-mail: lfluca01@louisville.edu

References

1. Tunis SR, Stryer DB, Clancy CM. Practical Clinical Trials. Increasing the value of clinical research for decision making in clinical and health policy. *JAMA* 2003; 290:1624-1632.
2. Wennberg J, Gittelsohn A. Small area variation in health care delivery. *Sci Am* 1973; 182:1102-1108.
3. Schuster MA, McGlynn EA, Brook RH. How good is the quality of health care in the United States? *Milbank Q* 1998; 76:517-

563.
4. Eddy DM, Billings J. The quality of medical evidence: Implications for quality of care. *Health Aff (Milwood)* 1988; 7:19-32.
5. McNeil BJ. Shattuck lecture – hidden barriers to improvement in the quality of care. *N Engl J Med* 2001; 345:1612-1620.
6. McQuay H, Moore A (eds). *An evidence-based resource for pain relief*. Oxford University Press, New York, 1998.
7. Hatala R, Guyatt G. Evaluating the teaching of evidence-based medicine. *JAMA* 2002; 288:1110-1112.
8. Manchikanti L, Singh V, Kloth D, 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.
9. 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.
10. American Pain Society. Guideline for the management of acute and chronic pain in sickle-cell disease. American Pain Society, Glenview, 1999.
11. American Geriatrics Society. The management of chronic pain in older persons: New guidelines from the American Geriatrics Society. *J Am Geriatr Soc* 1998; 46: 1998; 46: 128-150.
12. Sanders SH, Rucker KS, Anderson KO, Harden RN, Jackson KW, Vicente PJ, Gallagher RM. Clinical practice guidelines for chronic non-malignant pain syndrome patients. *J Back Musc Rehabil* 1995; 5:115-120.
13. 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.
14. Bogduk N. International Spinal Injection Society guidelines for the performance of spinal injection procedures. Part 1: Zygapophyseal joint blocks. *Clin J Pain* 1997; 13:285-302.
15. 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.
16. www.Cochrane.org.
17. Bigos SJ, Boyer OR, Braen GR, Brown K, Deyo R. Acute Low Back Problems in Adults. *Clinical Practice Guideline* Number 4. AHCPR Publication No. 95-0642. Agency for Health Care Policy and Research, Public Health Service, US Department of Health and Human Services, Rockville, December 1994.
18. Gonzalez EG, Materson RS (eds). *The Nonsurgical Management of Acute Low Back Pain*. Demos Vermane, New York, 1997.
19. Spitzer WO, Leblanc FE, Dupuis M (eds). Quebec Task Force on Spinal Disorders. Scientific approach to the assessment and management of activity-related spinal disorders: A monograph for clinicians. *Spine* 1987; (suppl) 12:1-59.
20. www.guideline.gov
21. 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.
22. Field MJ, Lohr KN. *Clinical Practice Guidelines: Directions for a New Program*. Washington DC, National Academy Press, 1990.
23. Steiner CA, Powe NR, Anderson GF, Das A. Technology coverage decisions by health care plans and considerations by medical directors. *Med Care* 1997; 35:472-489.
24. Garber AM. Evidence-based coverage policy. *Health Aff (Milwood)* 2001; 20:62-82.
25. Woolf SH, Hutchinson A, Eccles M, Eccles M, Grimshaw J. Clinical guidelines: Potential benefits, limitations, and harms of clinical guidelines. *BMJ* 1999; 318:527-530.
26. Grimshaw JM, Russell IT. Effect of clinical guidelines on medical practice: A systematic review of rigorous evaluation. *Lancet* 1993; 342:1317-1322.
27. Effective Health Care Bulletin (EHCB) Implementing clinical practice guidelines. Can guidelines be used to improve practice? *Effective Health Care Bulletin*, No. 6, University of Leeds, UK 1994.
28. Manchikanti L, Rivera J, Pampati V, Damron KS, Beyer CD, Barnhill RC. Effectiveness of clinical guidelines in interventional pain management. *Pain Physician* 2002; 5:127-132.
29. 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.
30. Shekelle P, Eccles MP, Grimshaw JM, Woolf SH. When should clinical guidelines be updated? *BMJ* 2001; 323:155-157.
31. 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.
32. 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. *JAMA* 2000; 283: 2008-2012.
33. Greenland S. Invited commentary: a critical look at some popular meta-analytic methods. *Am J Epidemiol* 1994; 140:290-296.
34. Lau J, Ioannidis JP, Schmid CH. Summing up evidence. *Lancet* 1998; 351:123-127.
35. Shapiro S. Meta-analysis/Shmeta-analysis. *Am J Epidemiol* 1994; 140:771-778.
36. Berlin JA. Invited commentary: benefits of heterogeneity in meta-analysis of data from epidemiologic studies. *Am J Epidemiol* 1995; 142:383-387.
37. Tunis R, Kang JL. Improvements in Medicare coverage of new technology. *Health Aff (Milwood)* 2001; 20:83-85.
38. Roland M, Torgerson DJ. What are pragmatic trials? *BMJ* 1998; 316:285.
39. Manchikanti L, Rivera JJ, Pampati V, Damron KS, MCManus CD, Brandon DE, Wilson SR. One day lumbar epidural adhesiolysis and hypertonic saline neurolysis in treatment of chronic low back pain: A randomized, double-blind trial. *Pain Physician* 2004; 7:177-186.
40. Manchikanti L, Rivera JJ, Pampati V, Damron KS, Beyer CD, Brandon DE, Wilson SR. Spinal endoscopic adhesiolysis in the management of chronic low back pain: A preliminary report of a randomized, double-blind trial. *Pain Physician* 2003; 6: 259-268.
41. Manchikanti L, Manchikanti KN, Damron KS, Pampati V. Effectiveness of cervical medial branch blocks in chronic neck pain: A prospective outcome study. *Pain Physician* 2004; 7:195-202.
42. Manchikanti L, Pampati V, Bakhit CE, Rivera JJ, Beyer CD, Damron KS, Barnhill RC. Effectiveness of lumbar facet joint nerve blocks in chronic low back pain: A randomized clinical trial. *Pain Physician* 2001; 4: 101-117.
43. Manchikanti L, Cash KA, Moss TL Rivera J, Pampati V. Risk of whole body radiation exposure and protective measures in fluoroscopically guided interventional techniques: a prospective evaluation. *BMC Anesthesiol* 2003; 3:2.
44. 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.
45. Manchikanti L, Pampati V, Rivera JJ, Beyer CD, Damron KS, Barnhill RC. Caudal epidural injections with Sarapin or steroids in chronic low back pain. *Pain Physician* 2001; 4:322-335.
46. Bogduk N, Karasek M. Two-year follow-up of a controlled trial of intradiscal electrothermal annuloplasty for chronic low back pain resulting from internal disc disruption. *Spine J* 2002; 2:343-350.
47. Lord SM, Barnsley L, Bogduk N. Percutaneous radiofrequency neurotomy in the treatment of cervical zygapophysial joint pain: a caution. *Neurosurgery* 1995; 35: 732-739.
48. Van Kleef M, Barendse GA, Kessels A, Voets HM, Weber WE, de Lange S. Randomized trial of radiofrequency lumbar

facet denervation for chronic low back pain. *Spine* 1999; 24:1937-1942.

49. Pauza KJ, Howell S, Dreyfuss P, Peloza JH, Dawson K, Bogduk N. A randomized, placebo-controlled trial of intradiscal electrothermal therapy for the treatment of discogenic low back pain. *Spine* J 2004; 4:27-35.
50. Carette S, Leclaire R, Marcoux S, Morin F, Blaise GA, St-Pierre A, Truchon R, Parent F, Levesque J, Bergeron V, Montminy P, Blanchette C. Epidural corticosteroid injections for sciatica due to herniated nucleus pulposus. *N Engl J Med* 1997; 336: 1634-1640.
51. Carette S, Marcoux S, Truchon R, Grondin C, Gagnon J, Allard Y, Latulippe M. A controlled trial of corticosteroid injections into facet joints for chronic low back pain. *N Engl J Med* 1991; 325:1002-1007.
52. Schwartz D, Lellouch J. Explanatory and pragmatic attitudes in therapeutic trials. *J Chronic Dis* 1967; 20:637-648.
53. Cherkin DC, Deyo RA, Battie M, Street J, Barlow W. A comparison of physical therapy, chiropractic manipulations, and provision of an educational booklet for the treatment of patients with low back pain. *N Engl J Med* 1998; 339:1021-1029.
54. Cherkin DC, Eisenberg D, Sherman KJ, Barlow W, Kaptchuk TJ, Street J, Deyo RA. Randomized trial comparing traditional Chinese medical acupuncture, therapeutic massage, and self-care education or chronic low back pain. *Arch Intern Med* 2001; 161:1081-1088.
55. A Medical Research Council. Streptomycin treatment of pulmonary tuberculosis. *BMJ* 1948; 2:769-782.
56. Byar DP, Simon RM, Friedewald WT, Schlesselman JJ, DeMets DL, Ellenberg JH, Gail MH, Ware JH. Randomized clinical trials: Perspectives on some recent ideas. *N Engl J Med* 1976; 1976; 295:74-80.
57. Feinstein AR. Current problems and future challenges in randomized clinical trials. *Circulation* 1984; 70:767-774.
58. Abel U, Koch A. The role of randomization in clinical studies: Myths and beliefs. *J Clin Epidemiol* 1999; 52:487-497.
59. Sacks H, Chalmers TC, Smith H Jr. Randomized versus historical controls for clinical trials. *Am J Med* 1982; 72:233-240.
60. Kjaergard LK, Als-Nielsen B. Association between competing interests and authors' conclusions: Epidemiological study of randomised clinical trials published in the BMJ. *BMJ* 2002; 325:1-4.
61. Djulbegovic B, Lacevic M, Cantor A, Fields KK, Bennett CL, Adams JR, Kuderer NM, Lyman GH. The uncertainty principle and industry sponsored research. *Lancet* 2000; 356:635-638.
62. Hopwood MB, Manning DC. Lumbar epidural steroid injections: Is a clinical trial necessary or appropriate? *Reg Anesth Pain Med* 1999; 24:5-7
63. Weinstein JN. The tortoise and the hare. Is there a place in spine surgery for randomized trials? *Spine* 1999; 24:2548-2549.
64. Fairbank J. Randomized controlled trials in the surgical management of spinal problems. *Spine* 1999; 24:2556-2561.
65. Carey TS. Randomized controlled trials in surgery. An essential component of scientific progress. *Spine* 1999; 23:2553-2555.
66. Lenrow DA, Chou LH. Randomized controlled trials in interventional spine: Perils and Pitfalls. *Pain Physician* 2003; 6:83-88.
67. Manchikanti L, Pampati V. Research designs in interventional pain management: Is randomization superior, desirable or essential? *Pain Physician* 2002; 5:275-284.
68. Kane RL. Approaching the outcomes question. In Kane RL (ed). *Understanding Health Care Outcomes Research*. Aspen Publication, Gaithersburg, 1997, pp 1-15.
69. Daly LE, Bourke GJ. Epidemiological and clinical research methods. In Daly LE, Bourke GJ (eds). *Interpretation and Uses of Medical Statistics*. Blackwell Science, 2000, Oxford, pp 143-201.
70. Ioannidis JPA, Haidich AB, Pappa M, Pantazis N, Kokori SI, Tektonidou MG, Contopoulos-Ioannidis DG, Lau J. Comparison of evidence of treatment effects in randomized and nonrandomized studies. *JAMA* 2001; 286:821-830.
71. Benson K, Hartz AJ. A comparison of observational studies and randomized, controlled trials. *N Engl J Med* 2000; 342: 1878-1886.
72. Concato J, Shah N, Horwitz RI. Randomized controlled trials, observational studies and the hierarchy of research designs. *N Engl J Med* 2000; 342:1887-1892.
73. West S, King V, Carey TS, Lohr KN, McKoy N, Sutton SF, 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.
74. 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.
75. Sutton AJ, Abrams KR, Jones DR, Sheldon TA, Song F. Systematic reviews of trials and other studies. *Health Technol Assess* 1998; 2: pp 1-276.
76. Cook DJ, Mulrow CD, Haynes RB. Systematic reviews: Synthesis of best evidence for clinical decisions. *Ann Intern Med* 1997; 126:376-380.
77. Lohr KN, Carey TS. Assessing 'best evidence': Issues in grading the quality of studies for systematic reviews. *Joint Commission J Qual Improvement* 1999; 25:470-*mission* 1999; 25:470- 479.
78. Mulrow CD. The medical review article: State of the science. *Ann Intern Med* 1987; 106:485-488.
79. Clark HD, Wells GA, Huet C, McAlister FA, Salmi LR, Fergusson D, Laupacis A. Assessing the quality of randomized trials: Reliability of the Jadad scale. *Control Clin Trials* 1999; 20:448-452.
80. Chalmers TC, Smith H Jr., Blackburn B, Silverman B, Schroeder B, Reitman D, Ambroz A. A method for assessing the quality of a randomized control trial. *Control Clin Trials* 1981; 2:31-49.
81. Hemminki E. Quality of reports of clinical trials submitted by the drug industry to the Finnish and Swedish control authorities. *Eur J Clin Pharmacol* 1981; 19:157-165.
82. Khan KS, Daya S, Jadad A. The importance of quality of primary studies in producing unbiased systematic reviews. *Arch Intern Med* 1996; 156:661-666.
83. Rotstein D, Laupacis A. Differences between systematic reviews and health technology assessments: A trade-off between the ideals of scientific rigor and the realities of policy making. *Int J Technol Assess Health Care* 2004; 20:177-183.
84. van Tulder M, Assendelft W, Koes B, Bouter LM. Method guidelines for systematic reviews in the Cochrane Collaboration Back Review Group for Spinal Disorders. *Spine*. 1997; 22:2323-2330.
85. 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.
86. Moher D, Schulz KF, Altman D, for the CONSORT Group. The CONSORT statement: revised recommendations for improving the quality of report of parallel-group randomized trials. *JAMA.* 2001; 285:1987-1991.
87. http://www.consort-statement.org/QUOROM.pdf
88. Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig LM, Lijmer JG, Moher D, Rennie D, de Vet HC; Standards for Reporting of Diagnostic Accuracy. Towards complete and accurate reporting of studies of diagnostic accuracy: The STARD initiative. *Clin Chem* 2003; 49:1-6. http://www.clinchem.org/cgi/content/full/49/1/1
89. Boswell MV, Singh V, Staats PS, Hirsch JA. Accuracy of precision diagnostic blocks in the diagnosis of chronic spinal pain of facet or zygapophysial joint origin. *Pain Physician* 2003; 6:449-456.
90. 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.
91. Shekelle PG, Andersson G, Bombardier C, Cherkin D, Deyo R, Keller R, Lee C, Liang M, Lipscomb B, Spratt K. A brief introduction to the critical reading of the clinical literature. *Spine* 1994; 19:2028S-2031s.