Selection of Workers' Compensation Treatment Guidelines: California Experience

Philip Harber, MD, MPH
Barbara O. Wynn, MA
Yee-Wei Lim, MD, PhD
Soeren Mattke, MD, PhD
Steven M. Asch, MD, MPH
Teryl K. Nuckols, MD, MSHS

Objective: *Workers' compensation systems increasingly use mandatory treatment guidelines to guide clinicians and for utilization management. This article describes the steps for selecting such guidelines.* ***Methods:*** *On the basis of experience with the RAND/University of California, Los Angeles project to help California select guidelines, we identified the necessary choices and processes for guideline selection and evaluation.* ***Results:*** *Major steps in guideline selection include: 1) define purpose; 2) assign decision-making authority; 3) decide whether to use existing guidelines or develop new ones; 4) choose whether to use one or multiple existing guidelines; 5) specify clinical topics that guidelines should address; 6) identify and screen guidelines; 7) evaluate guidelines; 8) consider implications of results; 9) select guideline(s); 10) disseminate selection; and 11) assess long-term effectiveness.* ***Conclusions:*** *Given the many choices required, selecting mandatory workers' compensation guidelines should involve careful planning and a transparent, well-defined process.* (J Occup Environ Med. *2008;50: 1282–1292*)

Workers' compensation (WC) costs have risen substantially over the past decade, particularly the proportion of total WC costs expended on medical care. In California, which has the largest state WC system in the United States, medical costs for injured workers grew by 111% between 1997 and 2002, when they came to represent more than half of total WC costs.[1] The cost increases were primarily driven by greater utilization rather than higher costs per service or other factors.[2]

In 2004, these increases prompted a critical evaluation of WC care in California. Throughout medicine, there has been growing reliance on explicit evidence of effectiveness, often embodied in clinical guidelines, to determine appropriate use of medical services. In this context, California legislators, seeking to control WC medical costs while maintaining quality, mandated the use of evidence-based treatment guidelines. Payers are now required to use these guidelines in utilization management (UM). The guidelines are also used in legal disputes over care, where they are considered "presumptively correct," that is the guidelines carry more weight than do the unsupported opinions of treating providers, although evidence can be presented to rebut the guidelines.[3–5] Thus, the State's choice of guidelines has enormous implications for injured workers, treating providers, and payers.

California legislators initially selected the American College of Occupational and Environmental Medicine (ACOEM) guidelines on an interim basis,[6] but simultaneously mandated a systematic evaluation of

From the Division of Occupational and Environmental Medicine, Department of Family Medicine (Dr Harber), David Geffen School of Medicine at the University of California, Los Angeles, Calif.; UCLA Center for Occupational and Environmental Health (Dr Harber), UCLA School of Public Health, Los Angeles, Calif.; The RAND Corporation (Ms Wynn, Dr Lim, Dr Mattke, Dr Asch, Dr Nuckols), Santa Monica, Calif.; VA Greater Los Angeles Healthcare System (Dr Asch), Los Angeles, Calif.; and Division of General Internal Medicine and Health Services Research (Dr Asch, Dr Nuckols), Department of Medicine, David Geffen School of Medicine at the University of California, Los Angeles, Calif.

This document includes content previously published in a RAND report, available on-line at: http://www.rand.org/publications/MG/MG400/.

Address correspondence to Philip Harber, MD, MPH, 10880 Wilshire Boulevard, Suite 1800, Los Angeles, CA 90024; E-mail: pharber@mednet.ucla.edu.

DOI: 10.1097/JOM.0b013e3181841fb4

TABLE 1
Five Guideline Sets Meeting All Screening Criteria

Guideline Name	Abbreviation	Developer	Last Review	Reference
Clinical Guidelines by the American Academy of Orthopedic Surgeons	AAOS	American Academy of Orthopedic Surgeons*	2001–2002	6
American College of Occupational and Environmental Medicine Occupational Medicine Practice Guidelines	ACOEM	American College of Occupational and Environmental Medicine*	2003	8
Optimal Treatment Guidelines, part of Intracorp Clinical Guidelines Tool(R)	IntraCorp	Intracorp†	2004	9
McKesson/Interqual Care Management Criteria and Clinical Evidence Summary	McKesson	McKesson Corporation†	2004	10
Official Disability Guidelines (ODG)—Treatment in Workers Comp	ODG	Work Loss Data Institute†	2004	11

*Medical specialty society developed guideline for use by treating providers.
†Corporation developed guideline for utilization management applications.

existing guidelines so that a more permanent choice could be made quickly.[3–5] The State contracted with the RAND Corporation (and its University of California, Los Angeles [UCLA] affiliates) to conduct a study that would develop a rational approach to identifying, screening, and evaluating existing treatment guidelines.

California is not alone in the process of selecting WC guidelines. Since 2004, several additional states have implemented mandatory treatment guidelines in their WC systems and others are considering doing so. There are many relevant guidelines from which to choose. Guidelines differ according to purpose, and certain selection criteria may be more appropriate for certain purposes. A systematic and transparent approach to guidelines selection should facilitate acceptance among WC stakeholders and may reduce potential legal challenges. Conversely, arbitrary selection or choosing inaccurate guidelines may lead to poor quality care or promote litigation.

In this article, we describe the guideline selection process used in California, highlighting the many considerations involved at each step. In addition to summarizing the approach used in California, we outline alternatives that were considered. We believe that future policymakers and others can use the study results to inform their attempts to develop a rational and publicly acceptable approach to selecting mandatory treatment guidelines for use in WC, and possibly other settings.

Materials and Methods

Identifying the steps of the California guideline selection process and the considerations entailed are the study's "results." Therefore, in the current section, we describe how these considerations were approached and provide an overview of the RAND study.

The study was a collaborative effort of the RAND/UCLA research team and California policymakers, including representatives of the California Division of Workers' Compensation (DWC, which monitors the administration of WC claims) and the California Commission on Health and Safety and Workers' Compensation (a joint labor-management body charged with overseeing the WC system and recommending administrative or legislative modifications). The RAND report to the State of California is available on-line http://www.rand.org/publications/MG/MG400/.[7]

The RAND/UCLA research team began the project by conducting a systematic search for guidelines addressing musculoskeletal disorders. Together with California's DWC and California Commission on Health and Safety and Workers' Compensation, we developed screening criteria for selecting promising guidelines identified in this search on the basis of guideline quality, comprehensiveness, and cost (see below). The criteria were publicly posted on the DWC Web site for comment. No comments were received, but many organizations submitted guidelines and other documents for consideration.

Applying the screening criteria left us with five comprehensive guideline sets (Table 1).[6,8–11] We used an internationally accepted tool (the Appraisal of Guidelines Research and Evaluation [AGREE] Instrument) to evaluate the technical quality of these guideline sets.[12–15] We then convened a multidisciplinary panel of national experts in the care of patients with work-related musculoskeletal disorders to assess the perceived comprehensiveness and validity of the five sets, adapting a modified-Delphi panel method to have panelists formally rate the guidelines.[16] We asked the panelists to focus on 10 selected clinical topics, any remaining material in each guideline set, and the overall guideline set. We synthesized the results of the AGREE and panel evaluations, presented them to a selected group of California stakeholders, and obtained initial feedback. We also conducted interviews with stake-

TABLE 2
Steps for Selecting Mandatory Workers' Compensation Treatment Guidelines

	Steps	Considerations
1	Define the purpose of the guidelines	Intended effects Mandate or voluntary implementation Needs of end users
2	Assign decision-making authority	Internal or external to government agency
3	Decide whether to use existing guidelines or develop new ones	Cost Timeframe Influence of local politics
4	Choose whether to use one or multiple existing guidelines	Comprehensive: one "set" from one developer selected to cover all conditions/interventions Patchwork: single guideline selected for each condition/intervention, with multiple developers Library: more than one guideline selected per condition/intervention
5	Specify clinical topics that guidelines should address	Conditions or interventions (tests or treatments) of interest Frequency, effects on health and cost For guideline sets, consider whether to examine all conditions/interventions of interest or a subset
6	Identify and screen guidelines	Guideline availability and cost Guideline quality
7	Evaluate guidelines	Evaluators: internal or external, expertise required Evaluation criteria: conceptual and practical Evaluation process: formal or informal
8	Consider implications of results	Unexpected results, including no adequate guidelines or several good guidelines Limitations of evaluation methods Practical considerations in implementing the selected guidelines Stakeholders' reactions
9	Select guideline(s)	Satisfy original mandate Attempt to resolve outstanding issues Future plans to resolve any limitations to selected guideline(s)
10	Disseminate selection	To end users
11	Assess long-term effects	Costs Access to care Quality of care Other effects of interest to stakeholders

holder groups around the state to identify issues raised during the interim implementation of the ACOEM guideline. Subsequently, RAND staff presented these findings at public forums. During the course of this study, many issues arose that had not been evident when the study started.

Results

Major steps in the California guidelines selection process are summarized in Table 1; some involved decisions by the California legislature and other policymakers rather than the RAND/UCLA team. Below, we discuss the considerations raised at each step, describe the range of possible approaches, and summarize the approach actually used in California. These are summarized in Table 2. Concrete examples are provided for each step; abbreviations for guideline sets are shown in Table 1.

Step 1: Define Purpose of the Guidelines. Create a Mandate: Regulatory or Voluntary?

Understanding why guidelines are being selected, ie, creating the mandate, is the first step and a critical one. It is particularly important for WC guidelines because, unlike the typical use of guidelines in medicine, WC care is governed by specific regulations that differ from state to state. Most practice guidelines are promulgated "for the general good," but guidelines in WC systems are often state specific and may impose constraints on the care provided. Furthermore, in some circumstances, guidelines can be more than "voluntary." Rather, they are implemented as a requirement—although they called "guidelines," they can actually be de facto legal regulations.

Several important questions should be answered during this first step. Are guidelines being selected because there is an explicit legal mandate, or was selection initiated by providers? Will implementation be voluntary or advisory? What purpose(s) will the guidelines serve? Who will be the principal end-user? For example, will providers use the guidelines to determine how to treat individual patients, will payers use them for UM, and will the judicial system refer to them during litigation?

Clinicians need somewhat different information from that needed by utilization reviewers, and guidelines are developed with these purposes in mind. Clinicians primarily need information about what they should do for individual patients; utilization reviewers primarily need information about what should not be done (or at least not paid for).[17,18] For example, the ACOEM guideline was developed to assist providers,[6] whereas several other guidelines are marketed to payers for use in UM.[9–11]

If guideline implementation is mandatory, or if the guidelines are given weight in litigation, then the stakes in guideline selection increase substantially. For example, if guidelines are considered "presumptively correct" in litigation, and treating providers' opinions carry less weight, then the specific guideline recommendations can profoundly affect a patient. No issues arise when appropriate care for an

individual patient is consistent with guideline recommendations. Nevertheless, appropriate care can be odds with guideline recommendations if the guideline is of poor quality or if an individual has unusual clinical characteristics. Providers will then find it challenging to ensure that appropriate care is provided because they must prepare for a legal challenge and assemble evidence to support their point. Thus, if guidelines are to be considered "presumptively correct" under the law, to the degree possible, their recommendations regarding appropriate care should be valid for most relevant patients.

CA Solution. Based on the CA legislation, the guidelines would be mandatory and emphasized use for utilization review. The state-selected guidelines must be used for UM and the same guidelines be "presumptively correct" in legal disputes. Some guidelines were much more UM oriented (eg, ODG) than others (eg, ACOEM). Nevertheless, this study sought the most valid guideline; none was excluded based on the stated purpose of the guideline developers (ie, both clinical and UM guidelines were evaluated).

Step 2: Assign Decision-Making Authority

Assigning decision-making authority means determining who is ultimately responsible for deciding which guideline will be implemented to satisfy the mandate. Because health and financial stakes are high for the various stakeholder groups, decision-making authority must be unambiguous. Both those with final decision-making authority and those who will provide formal input into the decision should be clearly identified. Even for nonmandatory situations, a detailed definition of the evaluative process before it begins will obviate accusations of changing the rules to favor certain guidelines.

Transparency in the overall guideline selection process and receptiveness to comments from all parties should engender a sense of fairness among stakeholders, even when their interests are considered but not ultimately satisfied. Nevertheless, when frank discussions among experts are needed, such as during evaluations of guideline clinical content, closed working sessions can shelter the experts from political influence. Similarly, a mechanism to limit undue influence from lobbyists, special interest groups, and product vendors is needed.

A second issue that requires clear decision-making authority is whether to conduct the process internally, or to contract for all or part of it. In some circumstances, governmental agencies transfer complete decision-making authority to an impartial outside organization. For example, in an analogous situation, states may mandate that impairments from work-related injuries be rated according to the American Medical Association Guides,[19] over which the state agencies have no direct control. In other instances, gathering and evaluation of information can be delegated to an outside organization, but the decision about how to act on the information is retained.

CA Solution. A hybrid approach was used, in which responsibility was shared between the government agency and an external contractor. The legislature gave ultimate responsibility for the decision to the Administrative Director of the Division of Workers' Compensation (DWC), a state agency. Nevertheless, the guideline evaluation process was externally contracted to the RAND/UCLA team, which was charged with making a specific recommendation to the DWC. Thus, although authority remained vested with the government agency, the selection process was largely entrusted to a contractor.

Step 3: Decide Whether to Use Existing Guidelines or Develop New Ones

A fundamental question is whether new guidelines should be developed explicitly for the WC system, or chosen from among existing guidelines? This is often termed, "Build it or Buy it?"

Rather than selecting from existing guidelines sets, it may be preferable to develop guidelines de novo. This helps to assure that the guidelines will reflect local practice patterns and available health care resources. It also increases the chances that the guideline format will be optimized to the intended purpose and ensures that the guidelines are readily available to the public, which may be challenging to guarantee if existing guidelines are used. Examples of states that have developed their own WC guidelines include Washington State and Colorado.[20,21]

Nevertheless, developing guidelines de novo also has several disadvantages. Guideline development requires considerable time and funding. Developing evidence-based guidelines entails an extensive literature review, and requires a systematic process by which experienced clinicians synthesize the literature and formulate recommendations about appropriate care. Because WC injuries include a wide range of ailments, development would require several expert committees. Finally, in a highly contentious arena such as WC, stakeholders may perceive state-specific guidelines to be "tainted" by local politics, creating a relative advantage for guidelines developed by out-of-state organizations.

CA Solution. CA elected to find existing guidelines rather than develop a new set. Several factors appear to have motivated the California legislature to use existing guideline sets rather than to develop new ones. Previously, the California WC system had developed state-specific guidelines under the aegis of the former Industrial Medical Council of the Department of Industrial Relations.[22] These guidelines were advisory rather than mandatory, and they had failed to restrain cost growth. In addition, there was not sufficient time to develop new guidelines, which can take years.

Step 4: Choose How to Cover Multiple Conditions: Single Comprehensive Set, Patchwork, or Library

Individual guidelines often focus on only one specific health condition or treatment, whereas WC systems need guidelines covering a wide variety of conditions and treatments. Thus, if one chooses to "buy rather than build," a basic decision about handling this multiplicity is needed. There are three major approaches:

i. Comprehensive set: This entails selecting a single "set" of guidelines from a particular developer. Each set contains many individual guidelines covering one or more conditions and treatments (eg, selecting guidelines from only a single source, such as ACOEM).
ii. Patchwork: In a patchwork approach, the single best available guideline is selected for each health condition or each treatment; the final product is a patchwork of component guidelines from different sources (eg, using ACOEM for shoulder problems but AAOS for lumbar problems).
iii. Library: In this approach, the WC system accepts multiple high-quality guidelines for each clinical topic (eg, allowing use of both the ACOEM and AAOS guidelines for back surgery, even if they partially conflict).

Each approach has advantages and disadvantages. The comprehensive set approach streamlines the evaluation process because fewer evaluations and selections are involved. It may produce greater consistency across conditions in terminology, general approach to evaluation and treatment, etc. Nevertheless, guideline quality may vary from topic to topic; in some cases, the guideline selected may not be the best guideline available.

The patchwork approach entails choosing the single best guideline for each topic. Nevertheless, a separate guideline evaluation process may be needed for each major organ system or intervention, imposing increased time and costs on the evaluation process. A patchwork of guidelines with few common elements may include inconsistent terminology or approaches, a problem that is particularly important when individual patients have multiple types of injuries. Inconsistencies may be quite disconcerting to end-users and patients; for example, a more aggressive attitude toward the use of surgery for one body part may conflict with a more conservative approach used for another body part.

The library approach best acknowledges the complexity of medical care by offering several different yet equally appropriate treatment approaches for each condition. Nevertheless, evaluation and dissemination may be more complicated than with other methods. Recommendations for some topics may conflict, creating uncertainty when the guidelines are used for UM or in legal disputes. Problems due to conflicting recommendations might be resolved by allowing care so long as one of the guidelines in the library supports its appropriateness. The library approach may be less acceptable to payers, but it might well be best for patient care and most acceptable to clinicians.

CA Solution. The California approach was to seek a single comprehensive set from a single developer if possible. This approach was chosen because there were several promising sets available and because a legislated deadline provided less than 6 months for an evaluation. (This proved to be only partially successful, as discussed below.)

Step 5: Specify the Clinical Topics That the Guidelines Should Address

The breadth of WC care precludes an in-depth evaluation of guidelines for every condition and intervention; it is necessary to focus on high priority issues to complete selection in a reasonable time frame. Furthermore, selection assumes that appropriate guidelines actually exist for the conditions and interventions considered; this may not be the case for less common conditions and interventions. We found that focusing our review on high priority issues required answering three questions. Which health conditions should be reviewed? Which interventions (ie, treatments or diagnostic procedures) should be reviewed? Should results of evaluations for one condition or intervention be extrapolated to other topics covered by guidelines from the same developer?

Possible criteria for choosing target conditions include frequency, severity of health effects, costs, and availability of data. The most common conditions are not necessarily those for which guidelines are needed. For example, minor skin irritation and simple lacerations are frequent but have only small effect on disability and health status, and they can usually be managed with acceptable outcomes. Health effects and costs may each be measured as the typical effect per worker with the condition or as the total effect across all workers with the condition. Possible health effect measures include self-reported health status, mortality, length of work disability, or residual permanent disability after reaching maximal medical improvement. Cost measures may be based solely on medical treatment costs or may include total cost for the case (including medical care, salary replacement indemnity, rehabilitation costs, and operational transaction costs). To be most comprehensive, costs should probably even include effects on worker productivity. The availability of data can influence the approach used. For example, financial data may be more available than data about frequency of disease, and assessing disability per se is challenging for a variety of reasons.

Other considerations are also relevant, including the availability of the intervention and the scope of the professional disciplines providing it. Some interventions have low per-unit costs, but in aggregate contribute greatly to costs per patient or overall. For example, chiropractic and physical therapy sessions are relatively inexpensive individually, but each individual patient may use many sessions. Conversely, certain surgical interventions are individually extremely expensive. The potential health effects—risks and potential benefits—should also be considered. For example, some types of surgery are associated with major complications. Focusing on treatment costs requires considering the most common treatments actually provided rather than those that are most traditionally accepted. For example, "alternative medicine" approaches such as chiropractic and acupuncture should be considered because patients with work-related musculoskeletal injuries turn to them frequently. The time course of care is another consideration. Some guidelines emphasize the early aspects of the care process, whereas others place greater emphasis on the more extensive therapy applied to individuals with advanced or nonresponsive disorders (eg, spinal surgery).

CA Solution. For this project, aggregate cost was the primary selection criterion. Emphasizing "cost drivers" would satisfy the mandate requiring guidelines that can be used to manage utilization, with the ultimate objective of curtailing costs by reducing inappropriate care. We identified interventions that were cost drivers in California WC (Table 3). Taken together, these interventions accounted for about 44% of the payments for professional services provided to California's injured workers, and associated surgeries accounted for about 40% of payments for inpatient hospital services.[7] The cost-driving interventions are typically used in the subacute, chronic, or late stages of treatment; interventions commonly used in early visits were not identified as significant cost drivers.

TABLE 3
Interventions Identified as Important Because They Were Cost-Drivers in California Workers' Compensation

MRI of the spine
Spinal injections
Spinal surgeries
Physical therapy
Chiropractic manipulation
Surgery for carpal tunnel and other nerve-compression syndromes
Shoulder surgery
Knee surgery

Scope and time frame issues prompted us to consider whether results should be extrapolated from one topic within a guideline to others. From the cost drivers in Table 3, we chose several condition/intervention pairs on which to focus our clinical evaluation: physical therapy, chiropractic manipulation and surgery for shoulder problems, carpal tunnel syndrome, and low back problems (with lumbar spinal decompression and fusion surgeries considered separately). For each guideline set, we assessed if the quality of the recommendations for these specific condition/therapy pairs could be used to reflect the quality of the guideline set as a whole. To test this, we had the expert panelists provide summary ratings for the remaining (residual) content within each guideline (ie, content other than for the selected condition/therapy pairs) and summary ratings for each guideline overall. Figure 1 shows the relationships between ratings for residual content and the average ratings for each of the specific items according to guideline set; this was done separately for validity and for comprehensiveness as described in more detail in Step 7 below. This empirical assessment supports the approach of generalizing from a limited number of assessments to describe the overall set.

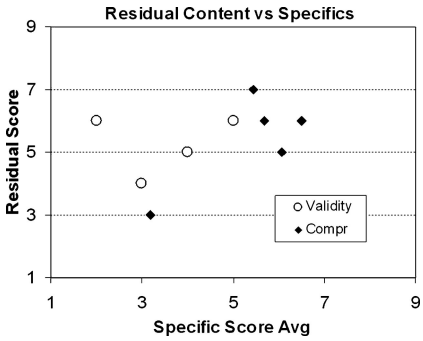

Fig. 1. Correlation of ratings for specific versus residual content. For each guideline set, the figure shows the average for all of the specific content items rated as well as the rating of residual content. Ratings for validity are shown by open circles and of comprehensives by filled diamonds.

Step 6: Identify and Screen Candidate Guidelines

A systematic guideline search strategy will prevent promising guidelines from being overlooked. Nevertheless, practical constraints preclude in-depth assessment of every guideline identified, so a screening process is necessary to identify guidelines that warrant detailed evaluation. Screening should be explicit and follow predetermined procedures to avoid concerns about possible bias. Subsequent steps evaluate the "best" among those meeting the screening criteria.

CA Solution. Because of the limited time for evaluation in California, the RAND/UCLA team applied screening criteria to eliminate all but the most promising guidelines. Criteria included (Table 4) 1) comprehensive guideline sets rather than guidelines addressing single topics (as explained above); 2) widely accepted standards for guideline quality; and 3) a cost ceiling for individual users of $500 (because the public might find greater costs to be prohibitive).

We identified guidelines from MEDLINE searches, National Guideline Clearinghouse lists, and internet searches. In addition, we contacted professional societies, state governments, and experts. This approach yielded 72 guidelines; the

TABLE 4
Screening Criteria for Guidelines Receiving Further Evaluation

Criterion	Definition	Source
A guideline	Systematically developed statements that assist practitioner and patient decisions about appropriate health care for specific clinical circumstances	Based on literature[36]
Evidence-based, peer-reviewed	Interpreted to mean based, at a minimum, on a systematic review of literature published in medical journals included in MEDLINE	Specified in legislation[3–5]
Nationally recognized	Interpreted to mean accepted by the National Guidelines Clearinghouse; published in a peer-reviewed U.S. medical journal; developed, endorsed, or disseminated by an organization based in two or more U.S. states; currently used by one or more U.S. state governments; or in wide use in two or more U.S. states	Specified in legislation[3–5]
Cover common and costly tests and therapies for disorders of spine, arm, and leg	Guidelines that covered several of the most common and costly tests and therapies in the California workers' compensation system: magnetic resonance imaging of the spine, spinal injections, spinal surgery, physical therapy, chiropractic manipulation, surgery for nerve compression syndromes, shoulder surgery, and knee surgery	Selected with CHSWC and DWC
Reviewed or updated at least every 3 yrs	Reviewed or updated between June 2001 and June 2004, and we confirmed a plan for updating at three year intervals or less	Based on literature,[37] selected with CHSWC and DWC
Developed by a multidisciplinary clinical team	A clinical team including at least three major types of providers that care for injured workers	Based on literature[12,36] selected with CHSWC and DWC
Cost less than $500 per individual user in California		Developed with CHSWC and DWC

screening criteria narrowed the list to five sets (Table 1).

Step 7: In-Depth Analysis of Guidelines Passing Screening: Identifying the "Best"

Choosing the guideline(s) that will ultimately be implemented should entail detailed analysis, with both evaluation criteria and an evaluation process that have been specified a priori. Most importantly, the evaluation should reflect the "mandate" and be consistent with the choices made in earlier steps. Ideally, the process used should be transparent and minimize the potential for obtaining biased results. Possible evaluation criteria include:

- Perceived expertise/status of developers: For example, some may consider guidelines promulgated by a highly reputed professional organization (eg, a physician specialty society) as best.
- Development process: Guidelines may be ranked according to the degree to which their development conformed to prescribed methods for writing guidelines. For example, the AGREE Instrument and National Guidelines Clearinghouse criteria define acceptable development methods.[12,23] The AGREE Instrument considers issues such as whether systematic methods were used to identify and grade the evidence, and formulate recommendations (see Table 5).[12]
- Format and presentation: The ease with which the intended end-users can easily navigate within the guidelines may be considered important. For example, clinicians may prefer a different format from utilization managers. Others may value summaries of the empirical studies supporting each guideline recommendation. The AGREE Instrument includes questions pertaining to clarity and presentation.[12]
- Content validity: Providers' confidence that the recommendations are valid may be important.
- Content comprehensiveness: Whether the guidelines pertain to most patients with the relevant conditions or undergoing the interventions identified in Step 5 may also be judged important.
- Acceptability to end-users: Is the candidate guideline acceptable to those who must ultimately use it, such as treating physicians or utilization managers?

Evaluation can be conducted in several ways. Evaluators can include staff of the organization conducting the evaluation, a committee of independent clinical experts, consultants, or even a panel of stakeholders having relevant expertise. They may use ad hoc methods, unstructured consensus methods, formal consensus techniques, or quantitative votes.

The natural confidence that providers place in the efficacy of their

TABLE 5
AGREE Instrument Domains

Domain and Definition
Scope and purpose: the overall aim of the guideline, the specific clinical questions and the target patient population
Stakeholder involvement: the extent to which the guideline represents the views of its intended users
Rigor of development: the process used to gather and synthesize the evidence, the methods to formulate the recommendations and to update them
Clarity and presentation: the language and format of the guideline
Applicability: the likely organizational, behavioral and cost implications of applying the guideline
Editorial independence: the independence of the recommendations and acknowledgement of possible conflict of interest from the guideline development group

From Ref. 12.

own specialty's interventions must be kept in mind when selecting a detailed evaluation approach. Treatment guidelines promulgated by one discipline may have a subtle bias favoring their approach and undervaluing others. Similarly, professionals participating in the guideline evaluation process have a vested interest in promoting their specialty and its interventions. For this reason, a multidisciplinary evaluation process is critical.

CA Solution. We used two assessment methods. First, RAND staff members used an explicitly defined algorithm (the AGREE Instrument) to the five selected guidelines (summarized in Table 1). This represented an efficient use of resources because researchers with relevant expertise were able to do these reviews with only a few weeks of effort. This contrasts with the much more extensive effort and participation of an expert panel process for the subsequent in-depth analyses.

In the second method, RAND adapted a relevant modified-Delphi panel method,[16] creating a new method in which expert clinicians rated and compared the five guidelines on comprehensiveness and validity. We incorporated a clinical panel evaluation because we were not confident that the AGREE evaluation methods would adequately distinguish among the guidelines, and we wanted to ensure that the guidelines were acceptable to providers with particular expertise in caring for occupational injuries.

We selected panelists so as to minimize potential conflicts of interest with evaluation outcomes. We included a range of specialties including occupational and environmental medicine, orthopedic surgery, neurosurgery, physical medicine, physical therapy, chiropractic, neurology, and internal medicine. We asked panelists to disclose any conflicts of interest, including whether they had strong a priori opinions against UM as a practice, and we permitted only 20% of panelists to be from California.

The mandate emphasizing guidelines for UM influenced our evaluation process. We asked panelists to rate the guidelines on how well they defined appropriate care for each of the 10 condition/therapy pairs above. Care that was inappropriate was, by definition, care that should not be provided and would be restricted by utilization managers after guideline implementation. For each topic, we asked panelists to rate how comprehensive were the recommendations pertaining to appropriateness. That is, did they apply to most patients with the condition who would be considered candidates for the therapy? We also asked them to rate the validity of the recommendations pertaining to appropriateness, meaning were they consistent with the panelists' perceptions of the literature or their expert opinions. Using the same approach, we asked panelists to rate how well the guidelines specified the quantity of care that should be provided, in those instances where the frequency, intensity, or duration of therapy can vary among patients. Given this project's time and resource constraints, we were unable to review the published research studies underlying each guideline. Nevertheless, the expert panelists considered any literature cited within the guidelines.

To illustrate the approach, examples of ratings of validity related to lumbar spine are shown in Fig. 2 for chiropractic, physical therapy, fusion surgery, and decompression surgery.

Step 8: Consider Implications of Evaluation Process Results

The implications of the evaluation findings must be considered carefully. Results may differ from what was expected, or reveal that no guidelines or several guidelines are adequate for implementation. The evaluation methods themselves may have important limitations. For example, our evaluation did not include systematic literature reviews. Stakeholders may have strong negative reactions about guidelines that do not satisfy their health or financial interests, and some guidelines may have practical limitations. For example, a guideline developed for UM may have a supporting electronic application tool, whereas a clinical guideline may not.

CA Solution. The two evaluation processes that we used produced somewhat incongruent results. The AGREE Instrument evaluation found all five guidelines to be of relatively high quality. In the evaluation by clinical experts, the ACOEM guideline was the highest ranked guideline, but panelists' ratings showed that perceived comprehensiveness and validity were not high for any of the five guideline sets examined and were highly variable for all of them as exemplified in Fig. 2. All of the guidelines were relatively weak for physical therapy, chiropractic manipulation, and lumbar spinal fusion; only AAOS was rated valid for this last topic. In addition, panel participants' comments revealed a general consensus that each of the five guidelines sets was barely satisfactory. Many felt that California should develop its own guidelines rather than using existing guidelines. The

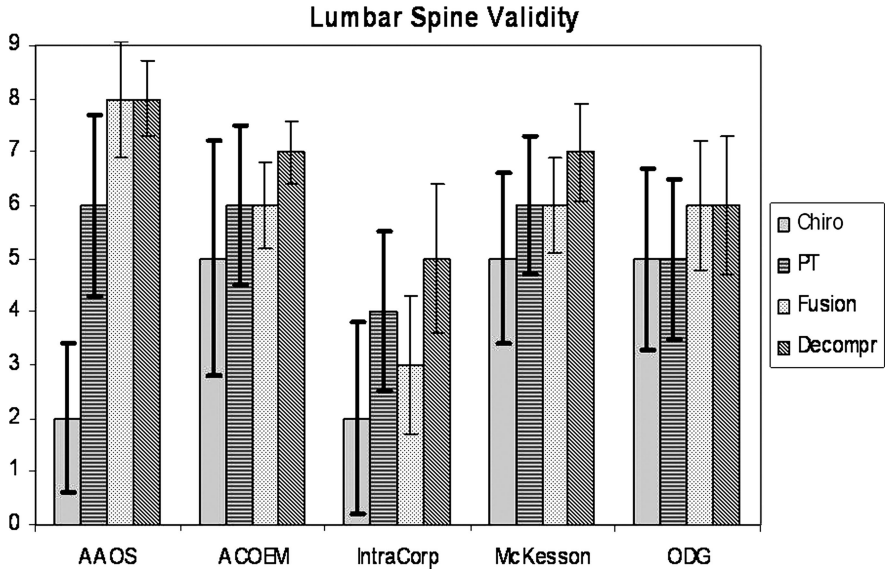

Fig. 2. Ratings of validity of lumbar spine guidelines. Ratings of the validity of each guideline's appropriateness recommendation are shown for four interventions for the lumbar spine: Chiropractic manipulation (solid gray), physical therapy (horizontal lines), fusion surgery (dots), and decompression surgery (diagonal lines).

RAND report describes evaluation results in detail.[7]

Selected stakeholders provided initial feedback on these results and their implications. They identified additional topics that they felt to be relatively weak in the ACOEM guideline, including chronic pain, acupuncture, and others. In general, they felt that the ACOEM guideline should be implemented as the principal guideline, but felt that switching to a "patchwork" approach, as described above, would be necessary to overcome the limitations in the ACOEM guideline.

Because the ACOEM guideline was already being implemented on an interim basis, we had an additional ability to explore the implications of continuing with this guideline. The RAND/UCLA team interviewed various groups of stakeholders around the state and learned that payers were having difficulty applying the ACOEM guideline consistently, possibly because it was not developed for UM purposes and lacked a tool supporting this application.

In our report to state policymakers, we recommended that California implement the ACOEM guideline as its overall guideline set, that the AAOS guideline be substituted solely for back surgery, and that better guidelines be identified for chronic pain, acupuncture, and certain other topics.

Step 9: Select Guideline(s) to be Implemented

In this culminating step, the individual or organization with decision-making authority selects the guideline(s) or guideline set(s) satisfying the original mandate and resolves as many other remaining issues as possible. In addition, this step should involve planning to address limitations of the selected guideline(s).

CA Solution. In California, the Administrative Director of DWC elected to implement the ACOEM guideline alone in 2005 and shortly after that AAOS withdrew its guideline. ACOEM subsequently developed a tool that supports use of its guideline for UM purposes.[24] Additional guidelines are now allowed to supplement the ACOEM guideline, and some major payers have been using ODG as a result (Phil Denniston, Work Loss Data Institute, personal communication, August 29, 2007). Recently, California policymakers proposed supplementary guidelines for chronic pain and acupuncture.

Step 10: Disseminate and Implement Selected Guideline(s)

In WC, guidelines are generally mandated for a specific purpose rather than simply published "for the general good." Even when guidelines are developed without a specific application in mind, dissemination is still important. For example, although the first edition of the ACOEM guidelines was published without any intent of direct regulatory application, there was a clear a priori plan for dissemination (publication by a publisher with close ties to the occupational health professional community).[6] When guidelines will have direct regulatory effects, there is increased need for effective dissemination to the multiple classes of users, which can include clinicians, insurers, WC judges, lawyers, patient representatives, etc.

CA Solution. In California, dissemination was conducted by state policymakers, representatives of ACOEM, and others. DWC provides public access to the ACOEM guideline at its local offices, the guidelines are available at many public libraries, and claims administrators are permitted to make copies for use in internal reviews.[25] Nevertheless, accessing guidelines on a public Web site, as done in states that have developed their own guidelines, might present fewer challenges to users.

Step 11: Assess Long-Term Effects

Ultimately, guideline implementation should be evaluated to determine whether the new policy has met the mandate defined in Step 1 and to identify corollary effects. For example, if cost reduction is a major goal, then after several years of actual utilization, it would be useful to evaluate whether medical costs were indeed reduced. Nevertheless, even if costs are the exclusive focus, it is also important to determine effects on other stakeholders, most importantly, on injured workers. Implementation of guidelines has the

potential to affect both access to WC care and the quality of that care.

CA Solution. In California, costs spent on WC medical care dropped between 2004 and 2005 and remained the same in 2005 and 2006.[26,27] Nevertheless, because several other reforms influencing cost were passed at the same time,[3–5] it is unclear the degree to which implementing guidelines is responsible for these trends. A systematic survey of providers and workers commissioned by DWC concluded that access to care was not a problem for most workers after the reforms. Effects on quality of care were not examined directly. Nevertheless, 64% of providers perceive UM to be a barrier to providing quality care and health outcomes are not very good; 1 year after their injuries occurred, 55% of workers have not fully recovered.[28]

Discussion

State WC systems are increasingly using treatment guidelines to limit the utilization of medical care and thereby curtail rising costs. When California recently implemented a mandatory guideline for this and related purposes, there was no existing protocol to use in the selection process. State policymakers and a RAND/UCLA research team found that many decisions arose in the course of choosing an acceptable guideline. In this report, we have described 11 critical steps and the considerations involved in each one. We found that formulating a systematic and transparent approach up front minimized the need to make ad hoc decisions later on, and seemed to produce results that stakeholders were more likely to accept.

We assumed that the characteristics of a good guideline are not specific to WC or UM; therefore, to assess the adequacy of guideline development, we used an evaluation method that has been applied in many other clinical contexts, the AGREE Instrument.[12–15] To assess the guideline content itself, we selected a well-established method previously used to develop tools for assessing the appropriateness of care, and adapted it to the purpose of evaluating treatment guidelines.[16] This afforded the advantages of both feasibility (it was not necessary to develop and validate new processes) and of credibility (because the technique has been validated for its original application). Consequently, however, we did not directly address certain aspects of care unique to WC, such as how well the guideline recommendations addressed the assessment of work ability status, work preclusions, and causation.

The legislated mandate to curtail inappropriate utilization led us to focus on conditions and interventions that are important drivers of medical costs, and such conditions and interventions generally pertain to later stages in the clinical course of work-related injuries rather than the first several visits. Nevertheless, a large randomized controlled trial has demonstrated that better care during the initial stages of musculoskeletal injuries, with careful attention to returning patients to work at the earliest appropriate opportunity, can substantially reduce temporary and permanent disability as well as total costs, yielding a return of $11.00 for each dollar invested.[29] Selecting only guidelines that can be used to curtail inappropriate utilization may, therefore, not be in the best long-term interests of either injured workers or WC payers; it may also be beneficial for states to select high-quality guidelines that describe optimal care early on.

Indeed, a much more comprehensive approach to improving the quality of the medical care that injured workers receive may be needed. Current outcomes for injured workers, at least in California, are far from optimal, with over half not having recovered after a year.[28] Prior studies examining quality of care across the US health care system as a whole have found it to be poor, with care that is consistent with rigorously developed quality measures only about half the time.[30] Similar studies examining the quality of WC care, particularly for the critical, early phases of treatment, have yet to be performed. Guidelines are one tool for improving quality, but providers often do not adhere to them, for a variety of reasons.[31–35] States can implement mandatory guidelines for UM relatively easily, but doing so for critical aspects of clinical care that are not routinely subject to utilization review would require a different oversight system, such as quality of care report cards or another mechanism. Given the possibility that improving quality might improve both clinical outcomes and reduce total costs, it may represent a unique opportunity to forge a consensus between injured workers and payers.

In summary, this study demonstrated that it is feasible to rationally select among guidelines even within the time and resource constraints that are common in the rapidly shifting world of state policy construction. The 11-step process outlined here should prove useful to other states that are considering implementing mandatory guidelines for use in UM and possibly to any policymaker considering guideline use for any condition. Policymakers should pay particular attention to the initial step in this process, developing the "mandate," meaning defining the intended purpose for which guidelines will be used. For example, in WC, implementing guidelines to limit the utilization of costly services may reduce medical expenditures, but the long-term need to improve the quality of medical care should also be considered because better care may improve both the physical health of workers and the fiscal health of WC payers.

Acknowledgments

Christine Baker, Executive Officer; Lachlan Taylor, Administrative Law Judge; and Irina Nemirovsky, Research Program Specialist II; of the California Commission on Health and Safety and Workers' Compensation; and Anne Searcy, Medical Director, of the California Division of Workers' Compensation participated in study conception and design.

Sponsored in part by the California Department of Industrial Relations (Division of Workers' Compensation and the California Commission on Health and Safety and Workers' Compensation) and the Southern California Education and Research Center (CDC/NIOSH).

References

1. California Workers' Compensation Institute. Frequently Asked Questions. Available at http://www.cwci.org/faq. Accessed November 11, 2004.
2. Telles CA. Workers Compensation Research Institute (Cambridge Mass.). *CompScope Benchmarks: Multistate Comparisons, 1994–1999.* Cambridge, Mass.: Workers Compensation Research Institute; 2004.
3. Alarcón. Workers' Compensation. State of California; 2003.
4. Calderon. Workers' Compensation: Administration and Benefits. State of California; 2002.
5. Poochigian. Workers' Compensation; 2004.
6. Glass LS. American College of Occupational and Environmental Medicine. Occupational Medicine Practice Guidelines: Evaluation and Management of Common Health Problems and Functional Recovery of Workers. Beverly Farms, Mass.: OEM Press; 2004.
7. Nuckols TK. Rand Corporation. *Evaluating Medical Treatment Guideline Sets for Injured Workers in California.* Santa Monica, CA: RAND Corporation; 2005.
8. American Academy of Orthopaedic Surgeons. Guidelines and Support Documents for Hip Pain, Knee Injury, Knee Osteoarthritis, Low Back Pain/Sciatica, Shoulder Pain and Wrist Pain. Rosemont, Illinois: American Academy of Orthopaedic Surgeons; 2001–2003.
9. Intracorp. *Intracorp Optimal Treatment Guidelines*. Philadelphia: Intracorp; 2003.
10. InterQual Care Management Criteria for Workers' Comp & Clinical Evidence Summaries. Newton, MA: McKesson Corporation; 2004.
11. Work Loss Data Institute. *Official Disability Guides (ODG), Treatment in Workers' Compensation*. San Diego: Work Loss Data Institute; 2004.
12. The AGREE Collaboration. The Appraisal of Guidelines for Research & Evaluation (AGREE) Instrument, 2001. London, UK: The AGREE Collaboration; 2001. Available at http://www.agreetrust.org/docs/AGREE_Instrument_English.pdf. Accessed March 21, 2007.
13. Burgers JS. Guideline quality and guideline content: are they related? *Clin Chem*. 2006;52:3–4.
14. 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.
15. Shiffman RN, Shekelle P, Overhage JM, Slutsky J, Grimshaw J, Deshpande AM. Standardized reporting of clinical practice guidelines: a proposal from the Conference on Guideline Standardization. *Ann Intern Med*. 2003;139:493–498.
16. Fitch K, Bernstein SJ, Aguilar MS, et al. *The RAND/UCLA Appropriateness Method User's Manual*. Santa Monica: RAND Corporation; 2001. MR-1269-DG-XII/RE.
17. Institute of Medicine (US). Committee on Utilization Management by Third Parties, Field MJ, Gray BH. *Controlling Costs and Changing Patient Care?: The Role of Utilization Management; Executive Summary.* Washington, D.C.: National Academy Press; 1989.
18. Wickizer TM, Lessler D. Utilization management: issues, effects, and future prospects. *Ann Rev Public Health.* 2002;23: 233–254.
19. American Medical Association. *Guides to the Evaluation of Permanent Impairment.* Chicago: AMA Press; 2000.
20. Washington State Department of Labor and Industries. Medical Treatment Guidelines: Guideline Process. Available at: http://www.lni.wa.gov/ClaimsIns/Files/OMD/MedTreat/GuidelineProcess.pdf. Accessed July 18, 2008.
21. Colorado Division of Workers' Compensation, Quality Improvement Program. Executive Summary of the Medical Treatment Guideline Case Review and Cost Studies. July 2000. Available at http://www.coworkforce.com/dwc/PUBS/execsummary.pdf. Accessed July 18, 2008.
22. California Department of Industrial Relations, Division of Workers' Compensation. Industrial Medical Council. Title 8 article 7. Practice Parameters for the Treatment of Common Industrial Injuries. Available at http://www.dir.ca.gov/dwc/MedicalUnit/IMC_ArchiveDocuments.html. Accessed July 18, 2008.
23. United States National Guideline Clearinghouse Inclusion Criteria. Available at http://www.guideline.gov/about/inclusion.aspx. Accessed July 6, 2006.
24. American College of Occupational and Environmental Medicine. Utilization Management Knowledgebase. Available at http://www.acoem.org/UMK.aspx. Accessed July 18, 2008.
25. California Department of Industrial Relations, Division of Workers' Compensation. Utilization Review. Access to the ACOEM Guidelines. Available at http://www.dir.ca.gov/dwc/UR_Main.htm. Accessed July 18, 2008.
26. Workers' Compensation Insurance Rating Bureau of California. 2005 California Workers' Compensation Losses and Expenses. Released: June 23, 2006. Available at https://wcirbonline.org/wcirb/wcirb_wire/pdf/2006/2005_loss_and_expenses.pdf. Accessed July 18, 2008.
27. Workers' Compensation Insurance Rating Bureau of California. 2005 California Workers' Compensation Losses and Expenses. Released: June 18, 2007. Available at https://wcirbonline.org/wcirb/resources/data_reports/pdf/2006_loss_and_expenses.pdf. Accessed July 18, 2008.
28. Kominski GF, Pourat N, Roby DH, Cameron ME. Access to Medical Treatment in the California Workers' Compensation System, 2006. Los Angeles, CA: UCLA Center for Health Policy Research; 2007. Available at http://www.healthpolicy.ucla.edu/pubs/publication.asp?pubID=216. Accessed July 18, 2008.
29. Abasolo L, Blanco M, Bachiller J, et al. A health system program to reduce work disability related to musculoskeletal disorders. *Ann Int Med*. 2005;143:404–414.
30. McGlynn EA, Asch SM, Adams J, et al. The quality of health care delivered to adults in the United States. *N Engl J Med*. 2003;348:2635–2645.
31. Bauer MS. A review of quantitative studies of adherence to mental health clinical practice guidelines. *Harv Rev Psychiatry*. 2002;10:138–153.
32. Cabana MD, Rand CS, Powe NR, et al. Why don't physicians follow clinical practice guidelines? A framework for improvement. *JAMA*. 1999;282:1458–1465.
33. Cruz-Correa M, Gross CP, Canto MI, et al. The impact of practice guidelines in the management of Barrett esophagus: a national prospective cohort study of physicians. *Arch Intern Med.* 2001;161:2588–2595.
34. Mosca L, Linfante AH, Benjamin EJ, et al. National study of physician awareness and adherence to cardiovascular disease prevention guidelines. *Circulation*. 2005; 111:499–510.
35. Switzer GE, Halm EA, Chang CC, Mittman BS, Walsh MB, Fine MJ. Physician awareness and self-reported use of local and national guidelines for community-acquired pneumonia. *J Gen Intern Med*. 2003;18:816–823.
36. Field MJ, Lohr KN. Institute of Medicine (US). Committee to Advise the Public Health Service on Clinical Practice Guidelines, United States. Dept. of Health and Human Services. *Clinical Practice Guidelines: Directions for a New Program.* Washington, D.C.: National Academy Press; 1990.
37. Shekelle PG, Ortiz E, Rhodes S, et al. Validity of the Agency for Healthcare Research and Quality clinical practice guidelines: how quickly do guidelines become outdated? *JAMA*. 2001;286: 1461–1467.