Eur Spine J (2006) 15: 543–553
DOI 10.1007/s00586-005-1027-y

REVIEW

Josep M. Arnau
Antoni Vallano
Anna Lopez
Ferran Pellisé
Maria J. Delgado
Nuria Prat

A critical review of guidelines for low back pain treatment

Received: 26 October 2004
Revised: 14 June 2005
Accepted: 21 August 2005
Published online: 11 October 2005

J. M. Arnau (✉) · A. Vallano · A. Lopez
Fundació Institut Català de Farmacologia, Hospital Universitari Vall d'Hebron, Universitat Autònoma de Barcelona, Passeig Vall d'Hebron 119–129, 08035 Barcelona, Spain
E-mail: jma@icf.uab.es
Tel.: +34-93-4283029
Fax: +34-93-4894109

J. M. Arnau · A. Vallano · A. Lopez
Servei de Farmacologia Clínica, Hospital Universitari Vall d'Hebron, Barcelona, Spain

J. M. Arnau · A. Vallano
Universitat Autònoma de Barcelona, Barcelona, Spain

F. Pellisé
Servei de Traumatologia i Ortopèdia, Hospital Universitari Vall d'Hebron, Barcelona, Spain

M. J. Delgado
Servei de Rehabilitació, DAP Sant Andreu, Barcelona, Spain

N. Prat
ABS Montornés-Montmeló, Barcelona, Spain

Abstract *Main problem*: Little is known about the methodological quality of guidelines for low back pain treatment. We evaluated the methods used by the developers according to established standards. *Methods*: PubMed, guideline databases, and the World Wide Web were used to identify guidelines. Seventeen guidelines met the inclusion criteria: interventions for low back pain stated, recommendations based on or explicitly linked to evidence, and English version available. Guidelines were evaluated independently by two appraisers using a practical tool for this purpose, *Users' Guides to the Medical Literature*, and the Appraisal of Guidelines for Research and Evaluation (AGREE) instrument. *Results*: Thirteen guidelines (76%) specified the most important therapies applied, but only nine (53%) included a complete description of the target population. Explicit processes to identify, select, and combine evidence were described in only six guidelines (35%). Few guidelines (3; 18%) explicitly considered all main outcomes when formulating therapeutic recommendations, and none contained a process to determine the relative value of different outcomes. Methodological criteria for grading the strength of the recommendations varied, and were often insufficiently specified. None of the guidelines assessed the impact of uncertainty associated with the evidence and values used. According to AGREE the quality score was highest for the scope and purpose, and clarity and presentation domains, and lowest for editorial independence and applicability. With regard to the recommendations, there was consensus for some of the interventions for acute pain (analgesics and NSAIDs, maintaining physical activity, and avoiding excessive bed rest), but explicit recommendations were lacking or ambiguous for 41% of the interventions. Most of the guidelines did not contemplate specific recommendations for chronic pain. *Conclusions*: A small number of the available guidelines for low back pain treatment achieved acceptable results for specific quality criteria. In general, the methods to develop the guidelines' therapeutic recommendations need to be more rigorous, more explicit and better explained. In addition, greater importance should be placed on the recommendations for chronic pain.

Keywords Low back pain · Guidelines · Practice guidelines · Quality assurance · Health care

Introduction

Low back pain is a common, growing problem and an important symptom-related reason for seeing a physician. The importance of low back pain in heath care is widely recognized and concerns have been raised on evidence of significant variation in the care provided, inappropriate and excessive use of diagnostic techniques and therapeutic interventions, and huge costs regarding its management [5, 16, 28].

Several clinical practice guidelines (CPG) have been developed in response to these unexplained variations in clinical practice and findings of inappropriate care and high costs [18, 32, 60]. However, favorable results with their application can be expected only if the guidelines have been properly designed and implemented [24, 29, 59]. The task of designing low back pain guidelines is particularly complex because of the large number and wide variety of therapeutic interventions used for this condition, and because developers are often faced with a limited number, if any, of adequate studies upon which to base the recommendations for some of these interventions.

Little is known about the appropriateness of the methods used by the developers of clinical guidelines for low back pain. The single, recently published, review of the quality of primary care guidelines for acute low back pain has shown disappointing results [54]. However, the therapeutic interventions reviewed, and the methods used by developers to assess and guarantee the validity and clinical relevance of the recommendations were not fully described in that study.

In 2001, the Catalan Agency for Health Technology Assessment and Research assigned special funds for the review of guidelines on therapeutic interventions for low back pain, in order to improve the methodological quality of guidelines for treating this condition. The purpose of this study was to systematically examine the currently available evidence-based CPG for low back pain and to describe and assess the methods used to create valid and clinically relevant recommendations regarding therapy for these patients, according to established standards.

Methods

Data sources

Clinical practice guidelines (CPG) were identified using specific search strategies in various sources:

a. MEDLINE and PubMed (from 1966 to June 2002). The search included combinations of the following keywords (MeSH terms): *low back pain* plus *guideline* or *practice guideline* or *clinical practice guideline*.
b. Guideline databases (up to June 2002), including the National Guidelines Clearinghouse, Scottish Intercollegiate Guidelines Group, New Zealand Guidelines Group, CMA Infobase, ACP-ASIM Guidelines, HSTAT, and ebm-guidelines. The search term used was *low back pain*.
c. World Wide Web (up to June 2002), by means of the browsers Google, Yahoo, and Altavista, and using the term *low back pain guideline*.
d. Additional guidelines were identified by manually searching the reference lists of retrieved reports and review articles [54].

Guideline authors were contacted to ask about unpublished guidelines or to request an English version of, or additional information on, published guidelines.

Guideline selection

Guidelines for low back pain had to meet the following criteria for inclusion in the study: (a) publication was within the period of 1992–2002 (when more than one version was available, only the most recent was included); (b) therapeutic interventions for low back pain were stated; (c) recommendations were evidence-based or evidence-linked (study references were included and/or specific grades of evidence were explicitly stated); (d) language was English or an English version was available.

Guidelines for low back pain were excluded on the basis of the following criteria: (a) focused on a single therapeutic intervention; (b) copy or summary of a previous guideline; (c) limited to a treatment algorithm without additional explanations; (d) narrative review without evidence-based recommendations; (e) focused exclusively on occupational health care; (f) objectives limited to teaching. The guidelines retrieved were checked for the inclusion criteria by three authors (JMA, AV, and AL).

Methodological evaluation and data extraction

All guidelines were reviewed independently by two authors (appraisers) and were scored for methodological quality according to the following publications: (a) the *How to Use* CPG (HUCPG) [29, 59] published in JAMA's Users' Guides to the Medical Literature series; (b) the Appraisal of Guidelines for Research and Evaluation (AGREE) instrument [49]. Assessment of recommendations only included those referring to therapeutic interventions.

The HUCPG contains ten guides clustered in three domains (Table 1). Each guideline included in the study was assessed according to the ten guides, and each guide

might contain one or more element to be appraised. After this independent review, the appraisers met and discussed their evaluations. When there was a discrepancy, the contents of the guideline were reviewed again to identify the sources of disagreement, and the evaluation was discussed again until consensus was reached. The main measure for the study was the percentage of guidelines that fulfilled the different elements in each guide.

For the present study, the most important pharmacological treatments were considered to be analgesics, NSAIDs and muscle relaxants and the most important non-pharmacological therapeutic interventions were education, physical activity or bed rest, exercise and spinal manipulation. Efficacy, risks, and cost were considered the most relevant outcomes of therapeutic interventions.

The recommendations made were classified into five categories: (1) the intervention is recommended in general, (2) the intervention is recommended in specific cases, (3) the intervention is discouraged in general, (4) the recommendation is ambiguous, and (5) a formal recommendation is lacking (although therapy might be reviewed in the guideline). Scales used to describe the strength of the recommendations and criteria defining the various categories in each scale were reviewed.

The AGREE is a validated generic instrument designed to assess the methodological quality of clinical guidelines [49]. It contains 23 items organized in six domains (Table 2). The main measure was the degree of quality expressed in percent term within a range of 0% (minimum) to 100% (maximum) for each of the six domains.

A descriptive statistical analysis was carried out using SPSS (Version 9.0).

Table 1 HUCPG guides

Are the recommendations valid?
Were all important options and outcomes clearly specified?
Was an explicit and sensible process used to identify, select, and combine evidence?
Was an explicit and sensible process used to consider the relative value of different outcomes?
Is the guideline likely to account for important recent developments?
Has the guideline been subjected to peer review and testing?
What are the recommendations?
Are practical, clinically important, recommendations made?
How strong are the recommendations?
What is the impact of uncertainty associated with the evidence and values used in the guidelines?
Will the recommendations help you in caring for your patients?
Is the primary objective of the guideline consistent with your objective?
Are the recommendations applicable to your patients?

Table 2 AGREE items

Scope and purpose
The overall objective(s) of the guideline is (are) specifically described
The clinical question(s) covered by the guideline is (are) specifically described
The patients to whom the guideline is meant to apply are specifically described
Stakeholder involvement
The guideline development group includes individuals from all the relevant professional groups
The patients' views and preferences have been sought
The target users of the guideline are clearly defined
The guideline has been piloted among the target users
Rigor of development
Systematic methods are used to search for the evidence
The criteria for selecting the evidence are clearly described
The methods used for formulating the recommendations are clearly described
The health benefits, side effects, and risks have been considered in formulating the recommendations
There is an explicit link between the recommendations and the supporting evidence
The guidelines has been externally reviewed by experts prior to its publication
A procedure for updating the guidelines is provided
Clarity and presentation
The recommendations are specific and unambiguous
The different options for management of the condition are clearly presented
Key recommendations are easily identifiable
The guideline is supported with tools for application
Applicability
The potential organizational barriers in applying the recommendations have been discussed
The potential cost implications of applying the recommendations have been considered
The guideline presents key review criteria for monitoring and/or audit purposes
Editorial independence
The guideline is editorially independent from funding body
Conflicts of interest of guideline development members have been recorded

Results

Descriptive characteristics

The search strategy identified 31 guidelines. A total of 14 (45%) were excluded [1, 3, 6, 11, 13, 33, 34, 39, 40, 43, 45, 48, 56, 58] because they were narrative reviews without evidence-linked recommendations (7) [1, 11, 13, 39, 43, 48, 58], they were copies or summaries of other guidelines (4) [3, 40, 45, 55], an English version was not available (2) [6, 33], or the focus was exclusively on occupational health care (1) [56]. The 17 guidelines included in the study [2, 8–10, 15, 17, 21, 26, 30, 31, 35–38, 41, 42, 50, 57] were published from 1994 to 2002 in the United States of America or Canada (9) [8, 21, 26, 30,

31, 35, 38, 41, 42, 50], European Union countries (5) [15, 17, 36, 37, 57], Australia and New Zealand (2) [2, 9], or Israel (1) [10]. Most of the guidelines were available on the World Wide Web (13; 76%) [2, 8, 9, 15, 21, 26, 30, 31, 35, 36, 41, 42, 50, 57]. Only three (18%) had been published in medical journals [10, 17, 41, 42]. Seven (41%) dealt with acute low back pain [2, 8, 9, 26, 30, 50, 57] and ten (59%) with both acute and chronic low back pain [10, 15, 17, 21, 31, 35–38, 41, 42]. In addition to therapeutic interventions, all but one [41, 42] guideline (94%) also recommended diagnostic and screening procedures. The intended setting for their use was defined in all but one [2] guideline (94%), and included primary care in all cases and hospital care in five (29%) [15, 21, 35, 37, 38].

Quality assessment according to the HUCPG

Validity of the recommendations Specification of all reasonable practice options and important outcomes Among a total of 64 therapeutic interventions for low back pain encountered, those included in more than half the guidelines are shown in Table 3. The most important pharmacological and non-pharmacological therapies (see methods) were specified in 13 (76%) guidelines. The outcomes (mainly efficacy and risks) of the therapeutic interventions included usually merited some general mention; however in seven guidelines (41%) the main outcomes defined by the developers were not specified. The efficacy, risks, and cost of the interventions were clearly specified in ten, five, and three guidelines, respectively. Only three (18%) guidelines clearly stated that all three outcomes had been considered by the developers.

Process used to identify, select, and combine evidence In most of the guidelines, a new review of evidence had been carried out; however, only seven guidelines (41%) explicitly described the methods used to identify evidence and seven (41%) the methods to select evidence. Both these processes were described in six guidelines. The main sources for identifying evidence were the Medline (7), Embase (4), and Cochrane Library (4) databases. The criteria for selecting evidence were most frequently related to study design, and were particularly addressed to selecting randomized clinical trials (7) or systematic reviews (6). In addition, most of the guidelines (14, 82%) included an explicit criterion (e.g. a scale) to combine available evidence and categorize it into different levels or grades. Explicit criteria to identify, select, and combine evidence were described in only six (35%) guidelines.

Process to consider the relative value of different outcomes None of the CPG included an explicit process to consider the relative value of the stated outcomes: efficacy, risks, and costs. Only six (35%) specifically described how decisions were taken (by consensus in all), and none explained how disagreement was solved. The members of the committee involved in the formulation

Table 3 Therapeutic recommendations for acute low back pain in guidelines[a]

Therapy	Recommendation, *n* (%)					Total guidelines, *n* (%)
	Use	Use in some cases	Do not use	Ambiguous	Absent	
NSAIDs	5 (45)	6 (55)	0	0	0	11 (100)
Non-opioid analgesics	11 (100)	0	0	0	0	11 (100)
Opioids	0	3 (33)	3 (33)	2 (22)	1 (11)	9 (100)
Muscle relaxants	1 (9)	1 (9)	6 (55)	3 (27)	0	11 (100)
Steroid injection (“trigger point”)	0	1 (14)	2 (28)	1 (14)	3 (44)	7 (100)
Steroid injection (epidural)	0	2 (28)	1 (14)	0	4 (58)	7 (100)
Physical activity	10 (91)	0	0	0	1 (9)	11 (100)
Bed rest	0	2 (17)	9 (75)	0	1 (8)	12 (100)
Exercises	2 (17)	5 (41)	0	2 (17)	3 (25)	12 (100)
Education	5 (56)	0	0	1 (11)	3 (33)	9 (100)
Manipulation	1 (9)	7 (64)	1 (9)	1 (9)	1 (9)	11 (100)
“Back school”	0	1 (10)	1 (10)	2 (20)	6 (60)	10 (100)
Transcutaneous electric nerve stimulation	0	0	3 (33)	1 (11)	5 (55)	9 (100)
Physical agents	1 (10)	1 (10)	0	4 (40)	4 (40)	10 (100)
Lumbar corsets	0	0	3 (33)	0	6 (66)	9 (100)
Acupuncture	0	0	4 (44)	0	5 (56)	9 (100)
Massage	0	1 (12)	0	1 (12)	6 (76)	8 (100)
Biofeedback	0	0	2 (29)	0	5 (71)	7 (100)
Traction	0	0	4 (50)	1(12)	3 (38)	8 (100)
Total	36 (20)	30 (17)	39 (22)	19 (10)	57 (31)	181 (100)

[a]References of guidelines included: [2, 8, 9, 15, 17, 26, 30, 36, 37, 41, 42, 50, 57]

and development of the guidelines were clearly specified in 14 (82%) cases. These were mainly general practitioners (12), orthopedic surgeons (11), and physiotherapists (10). Representatives of patients were included as members of the committees in five guidelines. Conflicts of interest among guideline development members were only recorded in one guideline.

Accountability for important recent developments Twelve (71%) guidelines included recent references (same year as guideline publication or the year before). In addition, seven (41%) guidelines mentioned a date for updating, but in five of them the date had passed when the guidelines were retrieved for the present study and a new edition was not yet available.

Peer review and testing Guidelines underwent external review by experts prior to their publication in five (29%) cases, and by potential users in three (18%) cases. However, only one (6%) guideline stated that it had been tested amongst its intended users and health outcomes had been assessed after its application.

Recommendations assessment Clarity and clinical importance of the recommendations Five, of the ten guidelines mentioning both acute and chronic low back pain, did not provide clearly differentiated recommendations for these two types of pain [10, 21, 31, 35, 38], and they were not included in the analysis of the therapeutic recommendations (Tables 3, 4). The most frequent recommendations specified in the guidelines for acute low back pain (12) [2, 8, 9, 15, 17, 26, 30, 36, 37, 41, 42, 50, 57] are shown in Table 3. Explicit recommendations for therapeutic interventions for acute low back pain were lacking or were considered ambiguous in 41% of the interventions. The most frequently specified recommendations for chronic low back pain in the five guidelines in which this could be assessed [15, 17, 36, 37, 41, 42] are shown in Table 4. Among the total, three of the guidelines [9, 26, 36] provided specific recommendations for lumbar pain with pain radiating to the leg, which mainly consisted of corticoid injections, NSAIDs and manipulation.

A systematic description of effect sizes on all outcomes (efficacy, risks, and cost) for each intervention was carried out in only one guideline (6%) [37]. A nonsystematic description of effect sizes, including only some interventions or outcomes, was made in six (35%). There was no description of effect sizes in ten (59%) guidelines.

Strength of recommendations Fourteen guidelines (82%) had a system for grading the scientific evidence and the strength of the recommendations. Twelve of these had only one scale for this purpose, whereas two graded evidence and strength of recommendations separately using two different scales. Three guidelines (18%) had no scales for this purpose. Table 5 shows methodological criteria for grading the evidence and strength of the recommendations. As can be seen, a specific definition as to how the criteria were assessed was frequently lacking. Other additional criteria used, but not shown in Table 5, were the number of studies available (11), the number of patients included in the studies (4), and the relevance of the studies, although methods for assessing these items were not specified. There was a median of four categories in the grading scales (minimum 2 and maximum 6). The higher grades were generally defined by higher quality or relevance of the studies, randomized clinical trial design and a substantial number of studies available. The lower grades were defined by a lack of evi-

Table 4 Therapeutic recommendations for chronic low back pain in guidelines[a]

Therapy	Recommendation, *n* (%)					Total guidelines, *n* (%)
	Use	Use in some cases	Do not use	Ambiguous	Absent	
NSAIDs	1	1	0	1	1	4 (100)
Non-opioid analgesics	2	1	0	1	0	4 (100)
Steroid injection (“trigger point”)	0	1	1	1	0	3 (100)
Steroid injection (epidural)	0	0	1	0	2	3 (100)
Antidepressants	0	1	1	0	1	3 (100)
Exercises	4	0	0	1	0	5 (100)
Manipulation	1	1	0	1	0	3 (100)
“Back school”	0	0	0	3	1	4 (100)
Transcutaneous electric nerve stimulation	0	1	1	1	1	4 (100)
Acupuncture	0	0	1	0	2	3 (100)
Massage	0	0	0	2	2	4 (100)
Traction	0	0	1	1	1	3 (100)
Total	8 (19)	6 (14)	6 (14)	12 (28)	11 (25)	43 (100)

[a]References of guidelines included: [15, 17, 36, 37, 41, 42]

Table 5 Criteria used for grading evidence and strength of recommendations

	Use of criteria in guidelines			Total[a], *n* (%)
	Not used, *n* (%)	Used but not defined, *n* (%)	Used and defined, *n* (%)	
Criteria related to evidence				
Susceptibility to bias of studies	6 (43)	0	8 (57)[b]	14 (100)
Quality of studies	1 (7)	9 (64)	4 (29)	14 (100)
How precise results are	13 (93)	1 (7)	0	14 (100)
Heterogeneity of results	8 (57)	6 (43)	0	14 (100)
Criteria related to outcomes				
Size of beneficial effect	8 (57)	6 (43)	0	14 (100)
Risks	11 (79)	2 (14)	1 (7)	14 (100)
Costs	12 (86)	2 (14)	0	14 (100)

[a]There was no scale for grading evidence or strength of recommendations in three guidelines
[b]The best study designs were usually considered to be randomized clinical trials

dence or evidence based only on expert (or panel) opinion. Overall, only four (24%) guidelines used pre-defined criteria to assess both the quality of the study and the susceptibility to bias of the study design.

Impact of uncertainty associated with the evidence and values used in the guidelines None of the guidelines considered the possibility that the effect size of a therapeutic intervention on the different outcomes could be higher or lower than their best estimate. None of the guidelines performed a sensitivity analysis to assess the impact of uncertainty on the estimates used as the basis for the recommendations.

In summary, only four guidelines [31, 37, 41, 42, 57] included a description of the size of the effects that the interventions had on the results as well as pre-defined criteria to assess the quality of the studies, and only one [37] provided this information systematically.

Usefulness of the recommendations for patient care Guideline objectives Guideline objectives were specifically stated in 14 (82%) cases, and they were usually focused on assisting physicians in clinical decision-making (11 guidelines included algorithms). Other, less frequent objectives were to enable evaluation of clinical practice (4) and cost-containment (4). Target users were clearly defined in 14 (82%) guidelines, and included general practitioners in all cases. The objectives and target users were explicitly described in 14 (82%) guidelines.

Applicability of recommendations A clear, explicit, complete description of the patient population to whom the therapeutic recommendations could be applied was given in nine (53%) guidelines. This information was incomplete or lacking in eight (47%) cases. The patients' views and preferences were considered in five (29%) guidelines.

To summarize, only six guidelines [8, 31, 36–38, 50] explicitly described the target users and gave a complete description of the target patient population in whom the recommendations should be applied.

Quality assessment with the AGREE appraisal instrument

Quality assessment of the guidelines according to the AGREE instrument is shown in Table 6. Higher scores were obtained for the scope and purpose domains (63%) and the clarity and presentation domains (53%). The editorial independence (22%) and applicability (21%) domains were associated with lower scores. Only five guidelines attained a score of 50% or more in the domain for *rigor of development* [8, 36, 37, 41, 42, 57]. The overall result of the appraisers'

Table 6 Quality assessment of guidelines with the AGREE appraisal instrument

Domain	Score median (%)	95% CI	Range (minimum score–maximum score)
Scope and purpose	63	40–86	28–100
Stakeholder involvement	38	15–61	17–75
Rigor of development	32	10–55	7–64
Clarity and presentation	53	29–77	25–87
Applicability	21	2–40	0–66
Editorial independence	22	2–42	0–42

assessment was to recommend the guidelines with provisos or alterations. In three cases the guidelines were considered unacceptable.

Discussion

Although significant efforts have been made to produce reliable CPG for treating low back pain, our main results show that the final product often has substantial flaws that make assessment of the guidelines difficult, or that raise questions as to the validity of the recommendations, their clinical importance and their usefulness in this specific health care situation. This is a source of concern, since these aspects are crucial for users of medical literature [29, 59] to decide whether a clinical practice guideline is suitable for incorporation as "standard of care."

It is widely recognized that CPG can be important tools for improving health care, and efforts to develop them are continuously increasing. However, as several authors have pointed out, the quality of many guidelines is dubious [19, 22, 47, 51]. There are deficiencies in many areas of guideline development, but the biggest problems reside in the identification, evaluation, and synthesis of the scientific evidence [44] that forms the basis for the recommendations and assures their validity for specific health care settings [23]. In addition, adherence to CPG is low [12], including those for low back pain [4, 20]. Although the physician's final decision to adopt a guideline is influenced by many factors, the perceived quality of the final product may be relevant.

Concern about the quality of guidelines is increasing, but the problem of poor quality is not easy to solve, since there are no well-defined, universally accepted standards for their development. In addition (and this is the case with low back pain guidelines), developers are often faced with a limited number, if any, of appropriately designed studies upon which to base the recommendations. Although some health care technology agencies and institutional bodies interested in promoting evidence-based health care have formulated rigorous developmental models, increasingly more practice guidelines are being created outside these initiatives. Many guidelines are not published in peer review journals and, apparently, many are not submitted to other external quality control processes. The need for systematic, rigorous evaluation of guideline quality is paramount, and several checklist approaches have been suggested to assess the methods used in their development and to formulate their recommendations [14, 25, 27, 29, 49, 59].

Assessment of the methodological quality of clinical guidelines for low back pain has deserved little attention, compared with the effort devoted to produce them. Recently, van Tulder et al. [54] performed a systematic review of primary care guidelines for acute low back pain using the AGREE instrument and concluded that the overall quality of the guidelines was very disappointing. The present study, performed almost simultaneously, supports the general conclusions of these authors, but it is focused more on the therapeutic recommendations and the methods used by the developers to assess and guarantee the validity and clinical relevance of the recommendations. Both studies provide complementary information, since the tools used for quality assessment and the methods to identify and select the guidelines for analysis were different. The present study did not include six of the guidelines present in the study by van Tulder, essentially because of the period of time studied [46], the language [6, 7, 33], and the exclusion of guidelines centered on a single intervention [17]. Eleven of the 17 guidelines analyzed by van Tulder et al. [54] are included in our study [2, 8, 10, 15, 17, 31, 36–38, 50, 57]. A summary of the comparison of the results obtained with the AGREE instrument in the 11 guidelines common to both studies is shown in Table 7. The results are similar except for the *clarity and presentation* domain, in which our scores are lower. In the domain for *rigor of development* the guidelines with scores of 50% or more are the same [8, 36, 37, 41, 42, 57].

To evaluate the quality of low back pain guidelines, we used the systematic approach proposed in the reference article of the *Users' Guides to the Medical Literature*

Table 7 A comparison of the results obtained using AGREE instrument[a]

Domain	Present study score		van Tulder et al. study score	
	Median	Range	Median	Range
Scope and purpose	72	44–100	89	56–100
Stakeholder involvement	46	22–75	50	33–83
Rigor of development	43	17–64	48	33–86
Clarity and presentation	54	25–87	83	63–100
Applicability	22	0–66	33	11–56
Editorial independence	25	0–42	33	0–50

[a]In the 11 guidelines common to both studies

series [29, 59]. The AGREE instrument [49], the only validated tool for this purpose, was added to achieve another perspective. We assumed the user's viewpoint when answering the main questions proposed in the HUCPG regarding the decision to adopt, adapt, or reject guideline recommendations. Even though some of the guidelines studied achieved acceptable results for specific HUCPG guides concerning the validity [8, 9, 15, 30, 50, 57], strength [31, 37, 41, 42, 57], and utility [8, 31, 36–38, 50] of the recommendations, none of them attained good results in the three areas, and we found important flaws in the majority, particularly for the first two domains.

We point out the lack of a specific method to determine the relative value of different outcomes and the very limited explanations on how decisions were taken and on external review processes. When the problems in handling the evidence and updating the guidelines are added to these factors, the overall impression of validity is poor. In general, the low back pain guidelines studied did not provide any information about the effect sizes of interventions on relevant outcomes. Moreover, in a considerable percentage of the interventions reviewed, the recommendation itself was lacking or ambiguous. In addition, the methods used to grade the evidence and the strength of the recommendations were often insufficiently specified, and none of the guidelines used a method that took into account the uncertainty associated with the evidence upon which recommendations were based. The overall impression is that the basis for judging the clinical importance and strength of the recommendations is weak.

The primary method for guideline assessment in this study, the proposed set of guides contained in the HUCPG, is a simple practical system that applies the methodology of evidenced-based medicine to the evaluation of CPG. The fact that it was designed for use by clinicians gives it an eminently pragmatic link to the reality of daily practice. A limitation of our study is the fact that the Users' Guides to the Medical Literature approach has not been previously applied or validated to systematically assess clinical guideline quality. To overcome, at least in part, this limitation, each guideline was reviewed by two appraisers to identify possible discrepancies and these were discussed until consensus assessment was obtained. In addition, we used the widely accepted, validated AGREE instrument, which also showed substantial weaknesses in most of the domains assessed. Another potential limitation was the exclusion of guidelines in languages other than English and the fact that the guidelines themselves may have left out crucial information on the processes used for their development. Both these factors are related to the feasibility of the study and we believe that they do not substantially affect the findings.

We wish to remark that CPG are a useful aid for assisting physicians in clinical decision-making, and an advance to better, evidence-based health care. Some of the guidelines for low back pain studied herein could contribute to these objectives, since they provide guarantees for the validity and strength of the recommendations (HUCPG) [8, 9, 15, 30, 50, 57] and rigor of development (AGREE) [8, 36, 37, 41, 42, 57]. Nevertheless, we found that, in general, the methods to develop them need to be improved and better explained. The effort and resources devoted to formulating guidelines could achieve better results with rather simple changes, such as the use of more explicit methods for identifying and selecting the evidence, improvements in several aspects of the development process related to the role of the various actors involved (including external reviewers), and better guideline implementation, as was recently suggested by van Tulder et al. [54]. In addition, we emphasize the critical importance of assessing and guaranteeing the validity and clinical relevance of the therapeutic recommendations. A well-established system should be used to grade evidence and recommendations, taking into account the susceptibility to bias of the available studies, and how precise and homogeneous the results are. It is essential to consider all the important therapeutic options and outcomes, to specify the processes used to consider the relative value assigned to each outcome, and to provide sufficient information on the expected effect sizes of each recommended intervention. Systematic reviews, such as the Cochrane reviews could be a useful source of information for these tasks.

With regard to the recommendations, there was a notable degree of consensus for some of the interventions for treating acute pain, including administration of analgesics and NSAIDs, maintaining physical activity and avoiding excessive bed rest. Nevertheless, explicit recommendations were lacking or ambiguous for a high proportion of the interventions and little importance was placed on chronic pain; in most cases there were no recommendations for chronic pain treatment or they were not clearly separated from those for acute pain. Future guidelines should place more emphasis on chronic pain and provide recommendations that are clearly differentiated from those formulated for acute pain.

When developing guidelines for clinical situations such as LBP, for which there are few adequately designed studies upon which the recommendations can be based, it is of utmost importance to follow well-defined and sufficiently explicit methods and to inform on the limitations of the available evidence. In this way the guarantees of validity and relevance can be evaluated. In the case of LBP it may be particularly difficult to generate evidence according to the generally accepted methodological criteria of high quality, for

example, the randomized, double-blind design of clinical trials may not be appropriate. However, specific recommendations on how research into therapeutic approaches for LBP should be undertaken are now available [52, 53] and they should be taken into consideration when assessing the efficacy of any interventions in the future.

As attempts to develop CPG increase, there is rising concern about the duplication of effort, suboptimal products, and wasting of resources. We believe that instead of developing new CPG with limited resources and expertise, it may be better to identify the most reliable available CPG, select the one that best fits local needs, and implement it properly. However, it is likely that physicians will continue to face practice guidelines of varying quality, and tools such as the HUCPG and AGREE instrument can be useful when deciding whether to adopt, adapt, or reject guideline recommendations.

Conclusion

A few guidelines for low back pain treatment achieved acceptable results for specific quality criteria and there was consensus for some of the interventions for acute pain. However, in general, the methods used to develop the therapeutic recommendations need to be more rigorous, more explicit, and better explained. In addition, future guidelines should place more emphasis on chronic pain and provide recommendations that are clearly differentiated from those formulated for acute pain.

Acknowledgments This study is supported by Grant 057/21/2001 from the Catalan Agency for Health Technology Assessment and Research. We thank Inma Fuentes, MD, for careful review of the study methods, and technical support, and Celine Cavallo for English language editing.

References

1. Abenhaim L, Rossignol M, Valat JP, et al. (2000) The role of activity in the therapeutic management of back pain Report of the international Paris task force on back pain. Spine 25:S1–S33
2. ACC, The National Health Committee (1997) New Zealand acute low back pain guide. Wellington, New Zealand; available at: http://www.nzgg.org.nz/library/gl_complete/backpain1/. accessed June 5, 2002
3. Agency for Health Care Administration and The Florida Department of Health (1999) University of Florida patients with low back pain or injury: Medical practice guidelines. Tallahassee, FL
4. Amstrong MP, McDonough S, Baxter GD (2003) Clinical guidelines versus clinical practice in the management of low back pain. Int J Clin Pract 57:9–13
5. Anderson GBJ (1999) Epidemiological features of chronic low-back pain. Lancet 354:581–585
6. Arzneimittelkommision der deutschen Artzeschaft (2000) Empfehlungen zur Therapie von Kreuzschmerzen. Köln, Germany
7. Bekkering GE, Hendriks HJM, Koes BW, et al. (2001) KNGF-richtlijn lage rugpijn. Ned Tijdschr Fysiother 111:1–24
8. Bigos S, Bowyer O, Braen G et al. (1994) Acute low back pain problems in adults: Clinical practice guidelines no. 14. AHCPR publications no. 95-0642. Rockville, MD: Agency for health care policy and research, public health service, us department of health and human services; available at: http://hstat.nlm.nih.gov/hq/Hquest/db/42/screen/DocTitle/odas/1/s/46074. Accessed August 30, 2002
9. Bogduk N (on behalf of The Australasian Faculty of Musculoskeletal Medicine) (1999) Evidence-based clinical guidelines for the management of acute low back pain. (Draft clinical practice guidelines for the management of acute low back pain.) The National Musculoskeletal Medicine. National Health & Medical Research Council. Available at: http://www.emia.com.au/MedicalProviders/EvidenceBasedMedicine/afmm/. Accessed June 12, 2002
10. Borkan J, Reis S, Werner S, Ribak J, Porath A (1996) Guidelines for treating low back pain in primary care The Israeli low back pain guideline group. Harefuah 130:145–151
11. Bratton RL (1999) Assessment and management of acute low back pain. Am Fam Physician 60:2299–2308
12. Cabana MD, Rand CS, Powe NR, et al. (1999) Why don't physicians follow clinical practice guidelines? A framework for improvement. JAMA 282:1458–1465
13. Cohen RI, Chopra P, Upshur C (2001) Low back pain, part 2 Guide to conservative, medical, and procedural therapies. Geriatrics 56:38–47
14. Cluzeau FA, Littlejohns P, Grimshaw JM, Feder G, Moran SE (1999) Development and application of a generic methodology to assess the quality of clinical guidelines. Int J Qual Health Care 11:21–28
15. Danish Institute for Health Technology Assessment (1999) Low back pain: frequency, management and prevention from HTA perspective. Dan Health Technol Assess 1(1)
16. Deyo RA, Weinstein JN (2001) Low back pain. N Engl J Med 344:363–370
17. Faas A, Chavannes AW, Koes BW, et al. (1996) NHG-Standaard 'Lage-Rugpijn'. Huisarts Wet 39:18–31
18. Field MJ, Lohr KM (1990) Clinical practice guidelines Directions for a new program. National Academy Press, Washington DC
19. Giannakakis IA, Haidich A-B, Contopoulos-Ioannidis DG, Papanikolaou GN, Baltogianni MS, Ioannidis JPA (2002) Citation of randomised evidence in support of guidelines of therapeutic and preventive interventions. J Clin Epidemiol 55:545–555
20. Gonzalez-Urzelai V, Palacio-Elua L, Lopez-De-Munain J (2003) Routine primary care management of acute low back pain: adherence to clinical guidelines. Eur Spine J 12:589–594

21. Gozna E (2001) For the workplace health safety and compensation commission of New Brunswick. Guidelines for the diagnosis and treatment of low back pain. WHSCC New Brunswick. Available at: http://www.grandroundsnow.com/whscc/whscc_pdfs/LBP_Guidelines_ENG.pdf. Accessed June 19, 2002
22. Graham ID, Beardall S, Carter AO, et al. (2001) What is the quality of drug therapy clinical guidelines in Canada? CMAJ 165:157–163
23. Graham RP, James PA, Cowan TM (2000) Are clinical practice guidelines valid for primary care? J Clin Epidemiol 53:949–954
24. Grimshaw JM, Russell IT (1993) Effect of clinical guidelines on medical practice: a systematic review of rigorous evaluations. Lancet 342:1317–1322
25. Hadorn DC, Baker D, Hodges JS, Hicks N (1996) Rating the quality of evidence for clinical practice guidelines. J Clin Epidemiol 49:749–754
26. Haig A, Standiford C, Alvarez D et al (1997) UMMC low back pain guidelines. University of Michigan Health System. Available at: http://www.med.umich.edu/1libr/guides/lowback.htm. Accessed June 12, 2002
27. Harbour R, Miller J (for the Scottish intercollegiate guidelines network grading review group) (2001) A new system for grading recommendations in evidence based guidelines. BMJ 323:334–346
28. Hart LG, Deyo RA, Cherkin DC (1995) Physician office visits for low back pain: frequency, clinical evaluation, and treatment patterns from a US national survey. Spine 20:11–19
29. Hayward RSA, Wilson MC, Tunis SR, Bass EB, Guyatt G (for the evidence-based medicine working group) (1995) Users' guides to the medical literature. VIII. How to use clinical practice guidelines. A. Are the recommendations valid? JAMA 274:570–574
30. Holbrook AM (2000) (Chair) for Ontario musculoskeletal therapy review panel Ontario treatment guidelines for osteoarthritis rheumatoid arthritis and acute musculoskeletal injury. Ontario program for optimal therapeutics. Toronto: Quenn's Printer of Ontario. Available at: http://www.ocpinfo.com/client/ocp/OCPHome.nsf/web/OPOT+Guidelines?OpenDocument. Accessed June 12, 2002
31. ICSI health care guidelines (2001) Adult low back pain. Institute for clinical system improvement Inc. Available at: http://www.icsi.org. Accessed October 14, 2002
32. Jackson MN, Nutting PA (1992) Clinical guidelines development: opportunities for family physicians. J Fam Pract 33:129–132
33. Keel P, Weber M, Roux E, et al. (1998) Kreuzxchmerzen: Hintergründe, Prävention Behandlung. Verbindung der Schweizer Arzte, Bern, Switzerland
34. Lothian health guidelines (2001) The management of acute low back pain. Lothian Health. Available at: http://www.nhslothian.scot.nhs.uk/primarycarelibrary/2_ClinicalPractice/2_Guidelines/Guidelines/AcuteLowBackPain.pdf Accessed August 25, 2002
35. Low back pain medical treatment guidelines (2001) State of Colorado. Department of Labor and Employment. Division of Workers' compensation. Available at: http://www.coworkforce.com/DWC/pdfrules/r17_exhibit_a_low_back_pain_with_links.pdf. Accessed November 15, 2002
36. Malmivaara A, Kotilainen E, Laasonen E et al. (2002) Clinical practice guidelines of the finnish medical association duodecim. Diseases of the low back. Available at: www.ebm-guidelines.com
37. Nachemson AL, Jonsson E (2000) Neck and back pain: The scientific evidence for causes, diagnosis, and treatment. Lippincott Williams & Wilkins, Philadelphia
38. NASS task force on clinical guidelines (2000) NASS clinical guidelines PHASE III—Unremitting low back pain. Version 1.0. North American spine society. Available at: http://www.spine.org/guidelines.cfm
39. Office of quality management. Clinical practice guidelines and disease management initiatives (2000) Acute low back pain. The cleveland clinic. Center for continuing education, USA. Available at: http://www.clevelandclinicmeded.com/ihpage2/acuteback.htm. Accessed June 12, 2002
40. Osteopath and Medicine working with NHS CSAG Report (1999) Acute low back pain guidelines. General osteopathic council, UK. Available at: http://www.osteopathy.org.uk/goc/medicine/guidelines.shtml/. Accessed June 12, 2002
41. Philadelphia panel (2001) Philadelphia panel evidence-based clinical practice guidelines on selected rehabilitation interventions: overview and methodology. Phys Ther 81:1629–1640. Available at: http://www.ptjournal.org/PTJournal/October2001/v81n10p1629.cfm. Accessed June 27, 2002
42. Philadelphia panel (2001) Philadelphia panel evidence-based clinical practice guidelines on selected rehabilitation interventions for low back pain. Phys Ther 8:1641–1674. Available at: http://www.ptjournal.org/PTJournal/October2001/v81n10p1641.cfm. Accessed June 27, 2002
43. Physician advisory committee (1996) Low back pain—treatment guidelines. Adopted by administrator of the Oklahoma workers' compensation court on January 2, 1996, USA. Available at: http://www.owcc.state.ok.us/2003%20Final%20Revisions%20-%20Low%20Back%20Pain.pdf/. Accessed June 12, 2002
44. Shaneyfelt TM, Mayo-Smiths MF, Rothwangl J (1999) Are guidelines following guidelines? The methodological quality of clinical practice guidelines in the peer-reviewed medical literature. JAMA 281:1900–1905
45. Spine algorithm task force (1996) Clinical guideline on low back pain. American academy of orthopaedic surgeons/North American spine society (AAOS/NASS), USA. Available at: http://www.guideline.gov/VIEWS/summary...mary&view=brief_summary&sSearch-string=-low+back+pain/. Accessed June 19, 2002
46. Spitzer WO, LeBlanc FE, Dupuis M, et al. (1987) Scientific approach to the assessment and management of activity-related spinal disorders. Spine 12:S1–S59
47. Sudlow M, Thomson RG (1997) Clinical guidelines: quantity without quality. Qual Health Care 6:60–61
48. The College of Physicians and Surgeons of Manitoba (1999) Guidelines and statements. Management of acute lower back pain, Canada. Available at: http://www.manitoba.ca/colleges/cps/Guidelines_and_Statements/343.html/. Accessed June 13, 2002
49. The Agree Collaboration (2003) Development and validation of an international appraisal instrument for assessing the quality of clinical practice guidelines: the AGREE project. Qual Saf Health Care 12:18–23
50. The Low Back Pain Workgroup (1999) VHA/DoD clinical guideline for the management of low back pain or sciatic in the primary care setting. Veterans Health Administration Department of Defense. Available at: http://www.oqd.med.va.gov/cpg/LBP/LBP_GOL.htm. Accessed November 11, 2002
51. Thomson R, McElroy H, Sudlow M (1998) Guidelines on anticoagulant treatment in atrial fibrillation in Great Britain: variation in content and implications for treatment. BMJ 316:509–513
52. van Tulder MW, Assendelft WJJ, Koes BW, Boutrer LM (editorial board of the cochrane collaboration back review group) (1997) Method guidelines for systematic reviews in the cochrane collaboration back review group for spinal disorders. Spine 22:2323–2330

53. van Tulder M, Furlan A, Bombardier C, Bouter L (editorial board of the cochrane collaboration back review group) (2003) Updated method guidelines for systematic reviews in the cochrane collaboration back review group. Spine 28:1290–1299
54. van Tulder MW, Tuut M, Pennick V, Bombardier C, Assendelft WJJ (2004) Quality of primary care guidelines for acute low back pain. Spine 29:E357–E362
55. Virtual Naval Hospital (1999) General medical officer (GMO) manual: clinical section: orthopedics. Low back pain. Department of the Navy. Bureau of Medicine and Surgery, USA. Available at: http://www.vnh.org/GMO/ClinicalSection/54lowbackpain.html/. Accessed June 13, 2002
56. Waddell G, Burton AK (2001) Occupational health guidelines for the management of low back pain at work: evidence review. Occup Med 51:124–135l. Available at: http://www.facoccmed.ac.uk/. Accessed June 6, 2002
57. Waddell G, McIntosh A, Hutchinson A, Feder G, Lewis M (1999) Low back pain evidence review. Clinical guidelines for the management of acute low back pain. Royal College of General Practitioners, London. Available at: http://www.rcgp.org.uk/rcgp/clinspec/guidelines/backpain/. Accessed April 10, 2002
58. Well-Connected (1998) WemMD—general guidelines for treatment of acute low back pain. WebMDHealth. Available at: wysiwyg://89/http://my.webmd.com/content/dmk/dmk_article_53207. Accessed June 13, 2002
59. Wilson MC, Hayward RSA, Tunis SR, Bass EB, Guyatt G (for the evidence-based medicine working group) (1995) Users' guides to the medical literature. VIII. How to use clinical practice guidelines. B. What are the recommendations and will they help you in caring for your patients? JAMA 274:1630–1632
60. Wolf SH (1990) Practice guidelines: a new reality in medicine I Recent developments. Arch Intern Med 150:1811–1818