Review

A critical appraisal of clinical practice guidelines for the treatment of lower-limb osteoarthritis

James N Pencharz*†, Elizabeth Grigoriadis‡, Gwenderlyn F Jansz§¶ and Claire Bombardier*†**††‡‡

*Institute of Medical Science, Faculty of Medicine, University of Toronto, Canada
†Arthritis & Autoimmunity Research Centre, University Health Network, Toronto, Canada
‡Department of Medicine, Sunnybrook and Women's College Health Sciences Centre, Toronto, Canada
§Department of Family and Community Medicine, St. Michael's Hospital, Toronto, Canada
¶Department of Family Medicine, Faculty of Medicine, University of Toronto, Canada
**Institute for Work & Health, Toronto, Canada
††Clinical Epidemiology and Health Care Research Program, Department of Medicine and Department of Health Administration, University of Toronto, Canada
‡‡Department of Medicine, Mt Sinai Hospital, Toronto, Canada

Correspondence: James N Pencharz, AARC-Health Care Research Division, 610 University Avenue, Suite 16-734, Toronto, ON, Canada, M5G 2M9. Tel: +1 416 946 4501 extension 4725; fax: +1 416 946 2291; e-mail pencharz@uhnres.utoronto.ca

Received: 5 July 2001
Accepted: 7 September 2001
Published: 16 October 2001

Arthritis Res 2002, **4**:36-44

(Print ISSN 1465-9905; Online ISSN 1465-9913)

Abstract

Clinical practice guidelines are important tools to assist clinical decision-making. Recently, several guidelines addressing the management of osteoarthritis (OA) have been published. Clinicians treating patients with OA must ensure that these guidelines are developed with consistency and methodological rigour. We undertook a qualitative summary and critical appraisal of six medical treatment guidelines for the management of lower-limb OA published in the medical literature within the past 5 years. A review of these six guidelines revealed that each possesses strengths and weakness. While most described the scope and intended patient populations, the guidelines varied considerably in the rigour of their development, coverage of implementation issues, and disclosure of conflicts of interest.

Keywords: clinical practice guidelines, osteoarthritis

Introduction

Clinical practice guidelines are "systematically developed statements to assist practitioner and patient decisions about appropriate health care for specific clinical circumstances" [1]. During the past 10 years, there has been a proliferation of clinical practice guidelines published in peer-reviewed literature. However, guidelines developed by different groups addressing the same clinical scenario have resulted in conflicting treatment recommendations [2]. Furthermore, recently completed structured evaluations of clinical practice guidelines led by Grilli [2] and Shaneyfelt [3] have raised concerns about the methodological quality of the guidelines' development.

Within the field of rheumatology, a number of clinical practice guidelines have recently been developed. In the year 2000, both the American College of Rheumatology (ACR) [4,5] and the European League Against Rheumatism (EULAR) [6] released guidelines for the medical management of lower-limb osteoarthritis (OA).

ACR = American College of Rheumatology; ADMMC = Algorithms for the Diagnosis and Management of Musculoskeletal Complaints; AGREE = Appraisal of Guidelines for Research & Evaluation; COX-2 = cyclooxygenase-2; EULAR = European League Against Rheumatism; ICSI = Institute for Clinical Systems Improvement; MeSH = Medical Subject Headings [of the National Library of Medicine]; NSAID = nonsteroidal anti-inflammatory drug; OA = osteoarthritis; PPIs = proton pump inhibitors.

OA is the most common form of arthritis [7]. It is a leading cause of chronic health problems and long-term disability and creates a significant economic burden for health services [8,9]. To date, however, the methodological quality of the various guidelines for the treatment of OA has not been formally evaluated and compared.

We performed a systematic literature search and evaluation of treatment guidelines for lower-limb OA. Our aim was to qualitatively document the recommendations provided within the guidelines, and quantitatively assess the methods used to develop them. The guidelines are intended for clinicians, who do not have the time to chase down additional information about their methodological development. We believe that guidelines should be reported in sufficient detail to allow clinicians to make judgements about the validity of their recommendations [10]. Therefore, we limited our evaluation of the guidelines to published materials or information easily accessible on the internet only.

Methods

The panel of four reviewers was comprised of an allied health professional and arthritis researcher (JP), a primary-care physician (GJ), a community rheumatologist (EG), and an academic rheumatologist and clinical epidemiologist (CB).

Our first step was to identify all relevant practice guidelines. To ensure the guidelines and recommendations were relevant and up to date, we evaluated only those published within the past 5 years. We searched OVID MEDLINE and OVID EMBASE using the following search strategy:

1. Osteoarthritis *or* osteoarthrosis *or* degenerative arthritis (Medical Subject Headings [MeSH] or keyword in title or abstract)
2. Practice guideline[s] *or* consensus development conference[s] (MeSH or publication type or keyword in title or abstract)
3. Standard[s] of Care *or* Practice Standard[s] *or* Practice Policy or Practice Parameter[s] *or* Practice Protocol[s] *or* Clinical Algorithm[s] *or* Practice Algorithm[s] *or* Guideline[s] *or* Recommendation[s]
4. 2 *or* 3
5. 1 *and* 4

The reference lists of all identified guideline documents were reviewed for other potential guidelines to be included in the evaluation.

Potential guidelines were reviewed for inclusion into the evaluation using specific criteria, based on Field and Lohr's [1] definition of practice guidelines. The guidelines chosen for review must have:

- been specific to the condition of OA. If it included other chronic pain or rheumatic conditions, then the discussion of OA must have been major focus.
- addressed the treatment of OA. Although some guidelines included clinical algorithms, a major focus of the guideline must have been treatment. It must have addressed either pharmacological or nonpharmacological therapy or both. Diagnostic and surgical guidelines were excluded. The guideline must have addressed a group or class of treatments (e.g. nonsteroidal anti-inflammatory drugs [NSAIDs] or opioid analgesics), not a single treatment (e.g. morphine).
- included a summary of the literature or evidence relating to the treatments discussed. The literature summary did not have to be a systematic one.
- been published in English.
- represented the opinions of at least five health professionals.

The MEDLINE and EMBASE searches identified 134 and 147 citations, respectively. Of these, 128 from MEDLINE and 142 from EMBASE were immediately excluded from consideration, because they were either non-English, editorials, letters to the editor, news stories or announcements, clinical studies, or clinical reviews. There remained 11 citations (6 from MEDLINE and 5 from EMBASE), of which we retrieved full text copies for review using the inclusion criteria described above.

Of the 11 potential documents [4–6,11–18], six met our inclusion criteria and were evaluated by the panel (Table 1). Five potential guidelines were excluded: four because they were clinical reviews and one because it was a duplicate of the EULAR guidelines. Each reviewer was provided with a package containing all the guidelines, background reading on the development and usage of guidelines, and copies of an instrument for the evaluation of guidelines, Appraisal of Guidelines for Research & Evaluation (the AGREE instrument).

This AGREE instrument was developed for use in assessing the quality of clinical guidelines [19]. It is the product of work coordinated by the AGREE Collaboration, an international panel of guideline experts charged with improving and standardizing all aspects of guideline development, implementation, and monitoring [20]. The instrument contains six themes, comprising 23 items, as well as an overall guideline global assessment of quality. The six themes are: scope and purpose, stakeholder involvement, rigour of development, clarity and presentation, applicability, and editorial independence. The guidelines' fulfilment of each item was rated on a 4-point response scale, from 4 = strongly agree to 1 = strongly disagree.

Each of the four reviewers independently evaluated the guidelines using AGREE. The panel then met and attained consensus on the majority of item ratings through informal discussion of each set of guidelines. The reviewers also discussed general strengths and weaknesses of the guidelines.

Table 1

Attributes of osteoarthritis treatment guidelines published within the past 5 years

Guideline [Reference no.]	Year published	Year literature search completed	Guideline development attributes	End users
Canadian Consensus Conference [11]	2000	Not stated	Rheumatologists; general and family practitioners Grading of evidence Formal consensus methods for recommendations External review	Primary-care physicians, rheumatologists
North of England [12,13]	1998	Not stated (Does state that recommendations cease to apply December 1999)	Multidisciplinary committee Meta-analyses of evidence Grading of evidence Informal consensus for recommendations Recommendation strength grading External review	Primary-care physicians
Algorithms for the Diagnosis and Management of Musculoskeletal Complaints (ADMMC) [14–16,33]	1997	Not stated	Multidisciplinary committee Grading of evidence Formal and informal consensus methods for recommendations Formal approval process by stakeholders	Primary-care physicians
ICSI [17,18]	1999 (2000)	Not stated	Multidisciplinary committee Grading of evidence Assume informal consensus for recommendations External review	Physicians, nurses, allied health professionals, health policy makers health care researchers
EULAR [6]	2000	December 1998	Rheumatologists, orthopaedic surgeons, and guideline methodologists Meta-analyses of evidence Grading of evidence Delphi consensus method Recommendation strength grading Internal review	Not stated
ACR [4,5]	2000	Not stated	Rheumatologists Informal grading of evidence Informal consensus for recommendations External review	Not stated

ACR = American College of Rheumatology; EULAR = European League Against Rheumatism; ICSI = Institute for Clinical Systems Improvement.

Results

Two of the six guidelines were restricted to the evaluation of NSAIDs: the Second Canadian Consensus Conference described evidence-based prescribing of NSAIDs for patients with OA or rheumatoid arthritis [11], while the North of England evidence-based guideline addressed the use of NSAIDs compared with basic analgesia in degenerative arthritis [12,13]. We also identified two comprehensive clinical algorithms: one entitled 'Algorithms for the Diagnosis and Management of Musculoskeletal Complaints' (ADMMC) [14–16] and the other, developed by the Institute for Clinical Systems Improvement (ICSI), focusing on the diagnosis and treatment of adult degenerative joint disease of the knee [17,18]. The other two guidelines were developed by professional bodies in rheumatology, the ACR [4,5] and the EULAR [6], and are comprehensive treatment guidelines addressing both pharmacological and nonpharmacological management.

The two guidelines restricted to the evaluation of nonsteroidal anti-inflammatory medication, the North of England Guidelines and the Canadian Consensus Guidelines, are described in the next two sections. The remaining four are then presented in a table.

North of England guideline

The North of England evidence-based guideline [12,13] compared the use of NSAIDs with basic analgesia (i.e. paracetamol) in treating the pain of degenerative arthritis.

The guideline addressed efficacy, gastrointestinal safety, and economic considerations. Its primary treatment recommendations were:

1. Paracetamol (acetaminophen) is the initial treatment of choice: up to doses of 4 g daily.
2. If paracetamol fails to provide symptomatic relief, then ibuprofen, at daily doses of 1.2 g, is the most appropriate alternative.

3. If further pain relief is needed, clinicians should consider increasing the dosage of ibuprofen to daily doses of 2.4 g, or using ibuprofen and paracetamol in combination, or using an alternative NSAID (i.e. diclofenac or naproxen), or initiating co-codamol (codeine).

The guideline discussed potential gastrointestinal side effects of NSAIDs. It encouraged clinicians to monitor patients closely for gastrointestinal toxicity and to review the need for chronic NSAID therapy regularly. Although the authors of the guideline acknowledged evidence supporting the use of misoprostol and proton pump inhibitors (PPIs) in preventing gastrointestinal injury, these therapies were not recommended routinely for all patients, with the possible exception of a selected group of high-risk patients (e.g. with a previous gastrointestinal bleed). Additionally, H_2 antagonists and PPIs may have a small impact on dyspepsia, but the benefits do not appear to exceed those from antacids.

From a cost–benefit perspective, the North of England guideline concluded that paracetamol and ibuprofen should be first- and second-line therapies, respectively, and that modified-release preparations are no more effective than standard treatment, while prophylactic use of misoprostol or PPIs is not cost effective.

Canadian Consensus Conference

The Canadian Consensus Conference [11] for the evidence-based use of NSAIDs in the treatment of OA and rheumatoid arthritis is a revised version of a NSAID guideline document first released in 1995. The guideline addressed issues of clinical efficacy, patient tolerability, gastrointestinal toxicity, renal toxicity, hypertension, and interactions with anticoagulants.

The primary recommendations were that NSAIDs should be the pharmacological agent of choice for patients with moderate to severe OA, while acetaminophen (paracetamol) can be considered as primary therapy in mild OA and as an adjunct therapy in moderate or severe OA. The authors of the guideline emphasized that decisions about the choice of pharmacological management should be made in concert with patients after discussing a drug's efficacy, safety, tolerability, and cost.

The Canadian Consensus Conference recommended cyclooxygenase-2 (COX-2) inhibitors as first-line therapy for patients with risk factors for a PUB (perforation, ulcer, bleed). The guideline stated that PPIs or misoprostol should be coadministered with any NSAID in patients with a history of an upper gastrointestinal bleed in the preceding 4 to 6 weeks. In patients with a history of an ulcer, testing for and eradicating *Helicobacter pylori* was recommended.

The Canadian Consensus Conference guideline grouped all NSAIDs (i.e. both nonspecific and COX-2-specific) as having similar effects on renal function and emphasized the need to monitor glomerular filtration rate (GFR). The guideline recommended that patients with hypertension have their blood pressure checked one week after being started on an NSAID and that a patient's antihypertensive medication dose or type may need to be modified appropriately.

Finally, this guideline indicated that COX-2-specific inhibitors should be selected over nonspecific NSAIDs in patients on anticoagulants (e.g. warfarin). Furthermore, a patient's international normalized ratio (INR) should be monitored frequently during the first week of NSAID use and warfarin doses should be modified appropriately.

ADMMC, ICSI, EULAR and ACR guidelines

Four of the guidelines we reviewed contained comprehensive recommendations for both pharmacological and nonpharmacological management of OA [4–6,14–18]. Their recommendations are summarized in Table 2. These guidelines indicated that nonpharmacological therapy such as patient education, social support, physical and occupational therapy, and exercise should be initiated as soon as possible and should represent the mainstay of therapy. Other nonpharmacological modalities recommended in some of the guidelines were weight loss, energy conservation, joint protection, heat, ice, acupuncture, massage, and electrical stimulation.

Similar to the North of England and Canadian Consensus guidelines, the four comprehensive guidelines recommend the use of acetaminophen (paracetamol) as the initial pharmacological therapy in mild OA. While the use of NSAIDs was suggested in all the guidelines if acetaminophen failed to control joint pain, the ACR guideline supported using NSAIDs as initial pharmacological therapy in patients with moderate to severe OA. COX-2-specific NSAIDs were recommended in the ACR guideline, and although they were not discussed in the published version [17], the ICSI guideline does recommend the use of COX-2-specific NSAIDs in its web update [18]. Because its literature search extended only up to December 1998, the EULAR guideline did not address the use of the COX-2 inhibitors. Conversely, the authors of the North of England guideline, which also excluded a discussion of COX-2-specific NSAIDs, acknowledged that information about new therapies would probably be available before their guideline's stated expiry date of December 1999.

Other pharmacological agents recommended as individual or adjunct therapies were opioid analgesics, corticosteroid or hyaluronic acid injections, and capsaicin cream. The ACR guideline provides the most comprehensive review of medication toxicity, including recommending using either misoprostol or PPIs as gastroprotective agents, whereas the ADMMC guideline only recom-

Table 2

Recommended therapies or modalities considered in four comprehensive treatment guidelines for lower-limb osteoarthritis published within the past 5 years[a]

	Guideline [reference]			
Therapy/modality	ADMMC OA [14–16] (Grade of evidence[b])	ICSI [17,18]	EULAR [6] (Strength of recommendation[c])	ACR [4,5]
Acetaminophen	Recommended (A)	Recommended	Recommended (A)	Recommended (mild to moderate OA)
NSAIDs	Recommended (A)	Recommended	Recommended (A)	Recommended (moderate to severe OA)
Cox-2-specific NSAIDs	Not discussed	Recommended (2nd-line therapy)	Not discussed	Recommended (2nd-line therapy in patient with high gastrointestinal risk)
Corticosteroid joint injection	Recommended (A)	Recommended	Recommended (A)	Recommended
Hyaluronic acid joint injection	Recommended (A)	Recommended	Recommended (B)	Recommended
Capsaicin cream	Recommended (A)	Recommended	Recommended (A)	Recommended
Other topical cream (e.g. NSAID cream)	Not discussed	Recommended	Recommended (A)	Recommended
Opioid analgesics	Recommended (B)	Recommended	Recommended (B)	Recommended
Aerobic exercise	Recommended (A)	Recommended	Recommended (A)	Recommended
Strengthening and range-of-motion exercise	Recommended (B)	Recommended	Recommended (A)	Recommended
Education	Recommended (no grade)	Recommended	Recommended (A)	Recommended
Arthritis self-management education	Recommended (A)	Recommended	Recommended (A)	Recommended
Physical therapy	Recommended (no grade)	Recommended	Recommended	Recommended
Occupational therapy	Recommended (C)	Recommended	Recommended	Recommended
Telephone contact	Recommended (A)	Recommended	Recommended (B)	Recommended
Weight loss	Recommended (B)	Recommended	Recommended (B)	Recommended
Walking assistive devices (e.g. shoe inserts, cane)	Recommended (C)	Recommended	Recommended (B)	Recommended
Assistive devices for activities of daily living	Recommended (C)	Recommended	Not discussed	Recommended
Glucosamine (with or without chondroitin)	Not discussed	Recommended	Recommended (A)	Not recommended
Knee brace or taping	Not discussed	Not discussed	Recommended (B)	Recommended
Hot pack or ice pack	Recommended (no grade)	Recommended	Not discussed	Recommended
Joint protection techniques	Recommended (no grade)	Recommended	Recommended (A)	Recommended
Planning activities	Recommended (no grade)	Recommended	Recommended (A)	Recommended
Acupuncture	Not discussed	Recommended	Not discussed	Not recommended
Electrical stimulation (e.g. TENS)	Not discussed	Recommended	Not discussed	Not discussed
Massage	Not discussed	Recommended	Not discussed	Not discussed

[a]This table covers only the four comprehensive treatment guidelines. For summary of recommendations of NSAID-specific guidelines, see Results section. [b]Grade of evidence: A = at least one high-quality randomized controlled trial with adequate power; B = evidence from underpowered randomized controlled; prospective controlled or historical controlled studies; C = consensus/expert opinion/uncontrolled trials; no grade = the therapy/modality was discussed in the text of the guideline, no level of evidence was provided to support the therapy/modality. [c]Strength of recommendation: A = directly based on category-1 evidence (meta-analysis of one or more randomized, controlled trials; B = directly based on category-2 evidence (at least one controlled study without randomization or at least one quasi-experimental study) or extrapolated recommendations from category-1 evidence; C = directly based on category-3 evidence (descriptive studies, such as comparative studies, correlation studies, or case–control studies) or extrapolated recommendations from category-1 or -2 evidence; D = directly based on category-4 evidence (expert committee reports or opinions and/or clinical experience of respected authorities) or extrapolated recommendations from category-2 or -3 evidence. ACR = American College of Rheumatology; ADMMC = Algorithms for the Diagnosis and Management of Musculoskeletal Complaints; COX-2 = cyclooxygenase-2; EULAR = European League Against Rheumatism; ICSI = Institute for Clinical Systems Improvement; NSAID = nonsteroidal anti-inflammatory drug; OA = osteoarthritis; TENS = transcutaneous electrical nerve stimulation.

mended the use of misoprostol as a gastroprotective agent. The EULAR and ICSI guidelines failed to describe an approach to gastroprotection.

Quantitative evaluation of the guidelines using the AGREE instrument

The quantitative evaluation of the guidelines using the AGREE instrument is summarized in Table 3. We scored each set of guidelines on each item of the AGREE instrument, using a 4-point scale ranging from 4 = strongly agree to 1 = strongly disagree. For reasons of clarity, we have dichotomized this scale. Guidelines were deemed to have fulfilled an item (indicated by + in Table 3) if the reviewers rated the guidelines as 3 or 4 on that item. Conversely, if the reviewers rated a guideline as 1 or 2 on any of the items then the item was deemed to have not been fulfilled (indicated by – in Table 3). In cases where one of the reviewers disagreed with the other three reviewers (i.e. rated the guideline as 3 or 4 when the other reviewers rated it as 1 or 2, or vice versa), the judgement of the majority of reviewers is reported in Table 3 and the opinion of the dissenting reviewer is indicated by the "b".

In general, the AGREE instrument was easy to apply to each of the guidelines and there were few disagreements between the reviewers. The scope and purpose for each guideline was usually clearly stated, but the stakeholder involvement in the development process was not always well described and varied considerably across guidelines. Although many of the guidelines were based on consultations with a range of health professionals caring for patients with arthritis, some tended to rely more on input from a narrow set of specialists [4–6] and only a few reflected formal consideration of patients' preferences [4,5,11]. The European guidelines [6,12,13] tended to be based on a more systematic approach to searching, selecting, and summarizing the evidence, but in all the guidelines, recommendations were explicitly linked to supporting evidence. In some, the strength of the supporting evidence was scored according to a predefined scale [6,12–18]. Only two contained results pooled across trials by the use of formal, quantitative meta-analysis tools and provided summary effect sizes for specific therapies [6,12,13]. In all six guidelines, recommendations were clearly highlighted and were supported at least by reference citations.

In contrast, most of the guidelines failed to adequately consider the important issues of dissemination, implementation, and applicability. The only exception was the ICSI guideline, which was specifically designed for implementation in a US Health Maintenance Organization [17,18]. In this guideline, efforts were put into the design of clinical decision tools and review criteria to monitor the success of its implementation. Finally, disclosure of financial support for the guideline process and independence from the funding body was poorly reported.

Discussion

Clinical practice guidelines will play an increasing role in clinical and health policy decision-making in the upcoming years. Consequently, the methodology used to develop guideline recommendations must be transparent and open to critical analysis, as in other areas of peer-reviewed research. Recent critiques of the guidelines have cast doubt on the recommendations brought forth by both the EULAR [21] and the ACR [22]. The scientific and clinical basis for those charges are outside the scope of this review.

We have made a qualitative summary (see Table 2) and a quantitative evaluation (see Table 3) of the content of treatment guidelines for lower-limb OA. The guidelines contained similar recommendations for both nonpharmacological and pharmacological therapies (see Table 2). This was particularly true in areas where there is strong evidence for the usefulness of certain nonpharmacological therapies (e.g. education, muscle strengthening, aerobic exercise, telephone contact) and pharmacological (e.g. acetaminophen, NSAIDs, opioids, hyaluronic acid injections). Conversely, in the case of therapies for which the evidence of usefulness is sparse or still accumulating (e.g. acupuncture, massage, glucosamine), either no recommendations were made or there was disagreement between guidelines.

In other fields, researchers have also found that guideline recommendations often diverged in areas of weaker evidence [23,24]. They believed that such disagreement could be explained by failure of the guideline developers to follow sound methodological principles in the formation of the guideline document [3]. We used the AGREE instrument, which incorporates six distinct themes of guideline development and implementation, to evaluate the methodological development of treatment guidelines for OA.

Scope and purpose

All the guidelines clearly stated their scope and purpose, which should help a prospective clinician identify the appropriate patient population to whom the guidelines should apply.

Stakeholder involvement

The authors of some of the guidelines elicited expert input only from a narrow group of specialists [4–6], which could lead to implementation difficulties [25]. This is particularly true when targeted health professionals, such as primary-care physicians or allied health professionals, were not included in the development process [2]. Furthermore, only the Canadian Consensus and ACR guidelines explicitly incorporated patient preferences into their recommendations.

Table 3

Ratings[a] according to the AGREE instrument of osteoarthritis treatment guidelines published within the past 5 years

AGREE instrument used for rating		Rating of guidelines [reference]					
Domain	Item	Canadian Consensus Conference [11]	North of England [12,13]	ADMMC OA [14–16]	ICSI [17,18]	EULAR [6]	ACR [4,5]
Scope and purpose							
	1. Overall objective(s)	+	+	+	+	+	+
	2. Clinical question(s)	+	+	+	+	+	–[b]
	3. Target patient population	+	+	+	+	+	+
Stakeholder involvement							
	4. Development group representative	+[b]	+	+	+	–	–
	5. Patient views and preferences	+	–	–[b]	–	–	+
Rigour of development							
	6. Systematic evidence search	–	+	–[b]	–	+	–
	7. Selection of evidence explicit	–	+	–	–	+	–
	8. Formulation of recommendations explicit	–	+	+	–	+	–
	9. Benefits, side effects, and risks described	+	+	+	+[b]	+	+
	10. Explicit link between evidence and recommendations	+[b]	+	+	+	+	+
	11. External review	+	+	+	–[b]	–	–
	12. Procedure for updating guideline	+	+	–	+	+	+
Clarity and presentation							
	13. Specific and unambiguous recommendations	+	+	+	+	+	+[b]
	14. Different treatment options	+	+	+	+	+	+
	15. Key recommendations easily identified	–[b]	+	+	+[b]	+	+[b]
Applicability							
	16. End users of guideline stated	+	+	+	+	–[b]	–
	17. Barriers to implementation are discussed	–	–	–[b]	+	–	–
	18. Cost implications are discussed	–	+	–	–	–	+
	19. Tools for application	–	–	–[b]	+	–	–
	20. Review/monitoring criteria defined	–	–	–	+	–	–
	21. Pilot testing	–	–	+	+	–	+
Editorial independence							
	22. Editorial independent from funding body	–	–	–	–	–	–
	23. Conflicts of interest are stated	–	+	–	–	–	+

[a]Ratings on a 4-point scale (4 = strongly agree, 1 = strongly disagree) have been dichotomized for purposes of clarity (see text): + indicates that the guideline has met the criterion; – indicates that it has not. [b]One of the raters disagreed (see text). ADMMC = Algorithms for the Diagnosis and Management of Musculoskeletal Complaints; ACR = American College of Rheumatology; EULAR = European League Against Rheumatism; ICSI = Institute for Clinical Systems Improvement.

Rigour of development

Within the guidelines that we evaluated, there existed shortcomings in the descriptions of the literature-search processes and how the evidence was subsequently selected for inclusion in the guideline. Three of the guidelines failed to appropriately describe how the literature was identified and selected, thereby raising the question of whether the literature review was comprehensive and unbiased in its coverage of the area of interest [26]. However, most of the OA guidelines clearly highlighted the link between evidence and recommendation by providing a table summarizing the evidence for each recommendation and, in some instances, the strength ratings of that evidence.

Although most guidelines called for external review from experts, stakeholders, and end users, a more explicit description of the external reviewers and the reason for their selection could have been provided.

To ensure that recommendations are relevant and timely, a method to monitor changes in the field and identify the need for modification of guideline recommendations should be implemented [27]. All but one guideline made reference to the need for revisions in the future, but they did not explicitly describe how this process would occur. This is particularly relevant given that two of the guidelines [6,14–16] failed to address the use of COX-2 inhibitors and two [6,17,18] failed to discuss gastroprotection during chronic long-term NSAID use. The failure of the guidelines to address these issues is due predominantly to the time frame of their literature searches.

Clarity and presentation

The key recommendations from most of the guidelines are clearly identifiable. Some of the guidelines embedded their recommendation with the text of a paragraph [4,5,11] located under a section heading, while others separately highlighted [6,12,13] their recommendations using italics at the beginning of a paragraph containing the discussion of that recommendation. We found the latter method a more effective technique for delivering guideline recommendations [28].

Applicability

Here we refer to the pilot testing and implementation barriers that must be overcome before the recommendations of a guideline are incorporated into everyday practice. In this respect, there is a clear weakness in the guidelines we reviewed, as almost none addressed implementation strategies. Implementation difficulties, however are often related more to political and organizational barriers than to scientific disagreements [29]. In addition, guidelines need to be adapted to local practice environments, because some treatment modalities may not be available locally (e.g. pool facilities, muscle strengthening equipment, COX-2-specific NSAIDs) [30]. All but the ICSI guidelines failed to address these difficulties adequately.

"Guidelines do not implement themselves" [1]. Continual monitoring of guideline implementation is essential. Developers can provide aids for clinicians and patients, such as a quick reference guide, patients' leaflets, or reminders in the patient chart, to help promote their use in clinical practice [31]. Furthermore, guideline developers should identify key review criteria or indicators that clinics or physicians should evaluate to determine the success of guideline implementation. The only guideline to fulfill this item was that of the ICSI.

Editorial independence

Guideline developers should fully disclose their funding sources, their editorial independence from their funding sources, and any other potential conflicts of interest. Full disclosure is necessary to ensure that any biases (or, more importantly, the perception of any biases) are acknowledged [22]. Although only two of the guidelines did not state their funding source, none of the guidelines completed all three of the disclosure tasks.

General comments

Because we have reviewed guidelines published in the peer-reviewed literature, we have captured a select group of OA treatment guidelines. Many guidelines developed by governmental or reimbursement agencies are never published in a peer-reviewed journal. To provide a review of these guidelines was beyond the scope of this appraisal. Paradoxically, however, guidelines developed locally are often the most effective at being accepted by end users and result in changes in professional practice [32].

Conclusions

We were interested in evaluating the methodological quality of treatment guidelines for lower-limb OA published in the peer-reviewed literature. Treatment recommendations agreed substantially between guidelines. However, the methodological quality of the guidelines varied and could generally be improved. Some of the guidelines failed to adequately incorporate the views of all stakeholders (e.g. patients, primary-care physicians, allied health professionals). We were particularly surprised by the failure of the authors of some of the guidelines to illustrate that their literature search, synthesis, and recommendation development processes were sufficiently rigorous. Furthermore, none of the guidelines addressed the 'doability' (i.e. implementation difficulties) of their recommendations. Future guideline developers in the field of rheumatology would benefit from transparent documentation of the process of their guideline development and its susceptibility to sources of bias.

Acknowledgements

This work was supported by the Arthritis & Autoimmunity Research Centre, University Health Network, Toronto, Canada.

References

1. Field M, Lohr K: G*uidelines for Clinical Practice: from Development To Use.* Washington, DC: National Academy Press, 1992.
2. Grilli R, Magrini N, Penna A, Mura G, Liberati A: **Practice guidelines developed by specialty societies: the need for critical appraisal**. *Lancet* 2000, **355**:103-106.
3. 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.
4. Altman RD, Hochberg MC, Moskowitz RW, Schnitzer TJ: **Recommendations for the medical management of osteoarthritis of the hip and knee: 2000 update**. *Arthritis Rheum* 2000, **43**:1905-1915.
5. American College of Rheumatology: **Guidelines for the development of practice guidelines** [http://www.rheumatology.org/research/guidelines/guidesonguides.html]. Last visited June, 2001.
6. Pendleton A, Arden N, Dougados M, Doherty M, Bannwarth B, Bijlsma JWJ, Cluzeau F, Cooper C, Dieppe PA, Gunther K-P, Hauselmann HJ, Herro-Beaumont G, Kaklamanis PM, Leeb B, Lequesne M, Lohmander S, Mazieres B, Mola E-M, Pavelka K, Serni U, Swoboda B, Verbruggen AA, Weseloh G, Zimmermann-Gorska I: **EULAR recommendations for the management of knee osteoarthritis: report of a task force of the Standing Committee for International Clinical Studies Including Therapeutic Trials (ESCISIT)**. *Ann Rheum Dis* 2000, **59**:936-944.
7. Lawrence RC, Helmick CG, Arnett FC, Deyo RA, Felson DT, Giannini ED, Heyse SP, Hirsh R, Hochberg MC, Hunder GC, Liang MH, Pillemer SR, Steen VD, Wolfe F: **Estimates of the prevalence of arthritis and selected musculoskeletal disorders in the United States**. *Arthritis Rheum* 1998, **41**:778-799.
8. Badley EM, Rasooly I, Webster GK: **Relative importance of musculoskeletal disorders as a cause of chronic health problems, disability, and health care utilization: findings from the 1990 Ontario Health Survey**. *J Rheumatol* 1994, **21**:505-514.
9. March LM, Bachmeier CJ: **Economics of osteoarthritis: a global perspective**. *Baillière's Clinical Rheumatology* 1997, **11**:817-834.
10. Woolf SH: **Practice guidelines: what the family physician should know**. *Am Fam Physician* 1995, **51**:1455-1463.
11. Tannenbaum H, Peloso PMJ, Russell AS, Marlow B: **An evidence-based approach to prescribing NSAIDs in the treatment of osteoarthritis and rheumatoid arthritis: the Second Canadian Consensus Conference**. *Can J Clin Pharmacol* 2000, **7(suppl)**:4A-16A.
12. Eccles M, Freemantle N, Mason J: **North of England evidence based guideline development project: summary guideline for non-steroidal anti-inflammatory drugs versus basic analgesia in treating the pain of degenerative arthritis**. *BMJ* 1998, **317**:526-530.
13. Eccles M, Freemantle N, Mason: **North of England evidence based guideline development project: methods of developing guidelines for efficient drug use in primary care**. *BMJ* 1998, **316**:1232-1235.
14. Lane NE, Thompson JM: **Management of osteoarthritis in the primary-care setting: an evidence-based approach to treatment**. *Am J Med* 1997, **103(6A)**:25S-30S.
15. Ellrodt AG, Cho Michaela, Cush JJ, Kavanaugh AF: **An evidence-based medicine approach to the diagnosis and management of musculoskeletal complaints**. *Am J Med* 1997, **103(6A)**:3S-6S.
16. **Algorithms for the diagnosis and management of musculoskeletal complaints**. *Am J Med* 1997, **103(6A)**:49S-80S.
17. Lee JA: **Adult degenerative joint disease of the knee: maximizing function and promoting joint health. Institute for Clinical Systems Integration**. *Postgrad Med* 1999, **105**:183-197.
18. Institute for Clinical Systems Improvement: **Degenerative Joint Disease: Health Care Guideline.** [http://www.icsi.org/guide/DJD.pdf]. Last visited June, 2001.
19. Cluzeau FA, Littlejohns P, Grimshaw JM, Feder G, Moran SE: **Development and application of a generic methodology to assess the quality of clinical guidelines**. *Int J Qual Health Care* 1999, **11**:21-28.
20. Appraisal of Guidelines for Research & Evaluation (AGREE) Collaboration [http://www.agreecollaboration.org/]. Last visited June, 2001.
21. Jawad AS: **EULAR recommendations for the management of knee osteoarthritis [letter]**. *Ann Rheum Dis* 2001, **60**:540.
22. Brandt KD, Bradley JD: **Should the initial drugs used to treat osteoarthritis pain be a nonsteroidal antiinflammatory drug?** *J Rheumatol* 2001, **28**:467-473.
23. Swales JD: **Guidelines on guidelines**. *J Hypertension* 1991, **11**:899-903.
24. Thomson R, McElroy H, Sudlow M: **Guidelines on anticoagulant treatment in atrial fibrillation in Great Britain: variation in content and implications for treatment**. *BMJ* 1998, **316**:509-513.
25. Feder G, Eccles M, Grol R, Griffiths C, Grimshaw J: **Clinical guidelines: using clinical guidelines**. *BMJ* 1999, **318**:728-730.
26. Woolf SH, Grol R, Hutchinson A, Eccles M, Grimshaw J: **Clinical guidelines: potential benefits, limitations, and harms of clinical guidelines**. *BMJ* 1999, **318**:527-530.
27. Browman GP: **Improving clinical practice guidelines for the 21st century: attitudinal barriers and not technology are the main challenges**. *Int J Technol Assess Health Care* 2000, **16**:959-968.
28. Jackson R, Feder G: **Guidelines for clinical guidelines**. *BMJ* 1998, **317**:427-428.
29. Hurwitz B: **Clinical guidelines: legal and political considerations of clinical practice guidelines**. *BMJ* 1999, **318**:661-664.
30. Ministry of Health and Longterm Care: *Ontario Drug Benefit Formulary I Comparative Drug Index*. Toronto: Queen's Printer for Ontario, 2001.
31. Grol R, Dalhuijsen J, Thomas S, in't Veld C, Rutten G, Mokkink H: **Attributes of clinical practice guidelines that influence use of guidelines in general practice: observational study**. *BMJ* 1998, **317**:858-861.
32. Gates PE: **Think globally, act locally: an approach to implementation of clinical practice guidelines**. *Journal of Quality Improvement* 1995, **21**:71-85.
33. Algorithms for the Diagnosis and Management of Musculoskeletal Complaints [http://www.swmed.edu/home_pages/cme/endurmat/lipsky/index.html]. Last visited June, 2001.