

REVIEW

EULAR recommendations for knee and hip osteoarthritis: a critique of the methodology

W Zhang, M Doherty

Br J Sports Med 2006;**40**:664–669. doi: 10.1136/bjsm.2004.016840

The quality of the EULAR recommendations for the management of hip and knee osteoarthritis (OA) was evaluated using a validated instrument. The quality and methods were compared with other guidelines and recommendations. EULAR recommendations were found to be among the best for overall quality. They show strengths with respect to scope, rigour of development, and clarity, but weaknesses with respect to stakeholder involvement, applicability, and editorial independence. However, a principal strength is their attempt to fill the gap between guidelines based solely on either research evidence or expert opinion. The methods used to synthesise research evidence (systematic review) and expert opinion (Delphi exercise) are robust. Strength of recommendation, based on combined consideration of research evidence, clinical expertise, and perceived patient preference, is valid and approaches the true essence of "evidence based practice" that considers each of these different forms of evidence.

between recommendations for knee and hip OA and subsequently developed a parallel and independent set of recommendations for hip OA.³ The rationale for this decision reflected the following multiple differences between hip and knee OA.

- Differences in anatomy and physiology: the hip is a very stable ball and socket joint with multiple planes of movement, whereas the knee is a modified hinge with one plane of movement and is prone to developing instability.
- Different risk factors for development of OA—for example, knee OA is female predominant, and obesity is an important risk factor, whereas hip OA is male predominant at younger ages but female predominant in older age, and obesity is only a modest risk factor.
- Differences in treatment applicability—for example, topical non-steroidal anti-inflammatory drugs (NSAIDs), mechanical bracing, and intra-articular injections are more suited to knee OA.
- Possible differences in response to the same treatment—for example, NSAIDs produce greater pain reduction in patients with knee OA than in those with hip OA⁹ and different effects from weight loss and diacerein.³

In recent years the European League Against Rheumatism (EULAR) Osteoarthritis (OA) Task Force has published separate recommendations for the management of knee OA (the first in 2000, with an update in 2003) and hip OA (2005).^{1–3} The methods used to develop these recommendations include expert consensus and systematic review of research evidence. Strength of recommendation for each specific key statement was also provided.

The recommendations have been widely disseminated in Europe and other parts of the world,^{4,5} and the methodology used by EULAR has been adapted by other guideline development teams.^{6,7} The EULAR recommendations for knee OA (2000) and the American College of Rheumatology (ACR) recommendations for both hip and knee OA (2000) were compared in 2003⁸ and reasonable agreement was found (table 1). In the same year as this comparison, EULAR published updated recommendations for knee OA with two major changes: (a) the order of the recommendations was optimised according to topic—that is, general, non-pharmacological, pharmacological, and surgical treatments (table 2); (b) new evidence was added to support the use of knee bracing, topical capsaicin, cyclooxygenase (COX)-2 selective inhibitors, gastro-protective agents, and opioids. However, unlike ACR, EULAR decided to maintain separation

In addition to the separation of knee from hip OA, there were other interesting and potentially important differences in recommendations that resulted from the different methodologies used by EULAR and ACR (table 3), despite both groups apparently examining the same research literature base. Thus methodology does appear to profoundly influence the recommendations that are generated. However, the quality of the EULAR recommendations in the context of all available OA guidelines remains unknown. In this article, we have undertaken a critical appraisal of the methodological strengths and weaknesses with respect to scope, stakeholder involvement, rigour, clarity, applicability, and editorial independence of the EULAR recommendations compared with other published guidelines. Such appraisal should prove useful in improving future guideline development and the quality of patient care.

Abbreviations: ACR, American College of Rheumatology; EULAR, European League Against Rheumatism; NSAID, non-steroidal anti-inflammatory drug; OA, osteoarthritis; RCT, randomised controlled trial; SOR, strength of recommendation; VAS, visual analogue scale

See end of article for authors' affiliations

Correspondence to:
Professor Doherty,
Academic Rheumatology,
University of Nottingham,
Clinical Sciences Building,
Nottingham City Hospital,
Nottingham NG5 1PB, UK;
michael.doherty@
nottingham.ac.uk

Accepted 1 June 2006

Table 1 Interventions evaluated by American College of Rheumatology (ACR) and European League Against Rheumatism (EULAR) guidelines

ACR and EULAR	ACR only	EULAR only
Non-pharmacological		
Patient education	Patellar taping	
Personalised social support	Occupational therapy	
Weight loss		
Aerobic exercise		
Muscle strengthening		
Range of motion exercise		
Walking aids		
Insoles		
Pharmacological		
Paracetamol	Topical capsaicin	Glucosamine/chondroitin
NSAIDs	COX-2 selective inhibitors	
IA corticosteroid	NSAIDs and gastroprotective drugs	
Topical NSAIDs	Analgesics	
IA hyaluronan		
Surgical		
Arthroplasty	Tidal irrigation	
	Arthroscopic debridement	
	Osteotomy	

Adapted with permission from Roddy and Doherty.⁸
NSAID, Non-steroidal anti-inflammatory drug; COX, cyclo-oxygenase; IA, intra-articular.

Table 2 Final set of 10 recommendations based on both evidence and expert opinion

- The optimal management of knee OA requires a combination of non-pharmacological and pharmacological treatment modalities
- The treatment of knee OA should be tailored according to:
 - Knee risk factors (obesity, adverse mechanical factors, physical activity)
 - General risk factors (age, comorbidity, polypharmacy)
 - Level of pain intensity and disability
 - Sign of inflammation—for example, effusion
 - Location and degree of structure damage
- Non-pharmacological treatment of knee OA should include regular education, exercise, appliances (sticks, insoles, knee bracing), and weight reduction
- Paracetamol is the oral analgesic to try first and, if successful, the preferred long term oral analgesic
- Topical applications (NSAIDs, capsaicin) have clinical efficacy and are safe
- NSAIDs should be considered in patients unresponsive to paracetamol. In patients with an increased gastrointestinal risk, non-selective NSAIDs and effective gastroprotective agents, or selective COX-2 inhibitors should be used
- Opioid analgesics, with or without paracetamol, are useful alternatives in patients in whom NSAIDs, including COX-2 selective inhibitors, are contraindicated, ineffective, and/or poorly tolerated
- SYSADOA (glucosamine sulphate, chondroitin sulphate, ASU, diacerein, hyaluronic acid) have symptomatic effects and may modify structure
- Intra-articular injection of long acting corticosteroid is indicated for flare of knee pain, especially if accompanied by effusion
- Joint replacement has to be considered in patients with radiographic evidence of knee OA who have refractory pain and disability

Adapted with permission from Jordan *et al.*¹
OA, Osteoarthritis; NSAID, non-steroidal anti-inflammatory drug; COX, cyclo-oxygenase; ASU, avocado soybean unsaponifiable.

Table 3 Differences in methodology between American College of Rheumatology (ACR) and European League Against Rheumatism (EULAR) recommendations published in 2000

Element	ACR	EULAR knee OA
Committee composition	4 American rheumatologists	18 rheumatologists, 3 orthopaedic surgeons, 2 evidence based medicine experts from 12 European countries
Management topic	Knee and hip OA	Knee OA
Literature review	Not described, up to 2000	Systematic review, up to 1998
Language of publication	Not defined	All European
Appraisal of literature	Principles of evidence based medicine	Quality scores, effect size, number needed to treat, category of evidence, strength of recommendation
Number of recommendations	Not specified	Top 10
Consensus method	Not described	Delphi exercise
Future research agenda	Not specified	Agreed using Delphi exercise

Adapted with permission from Roddy and Doherty.⁸

METHODS

The EULAR recommendations for hip OA and the latest version for knee OA formed the basis of this critical appraisal.¹⁻³ Guidelines/recommendations on the same topic were searched systematically using Medline, EMBASE, CIHINAL, AMED, WOS, and Google. Guidelines were included if they were (a) specifically for the management of hip and/or knee OA, and (b) systematically developed by a group of experts with either a consensus approach and/or an evidence based approach. If the guidelines had been updated, the most recent version was included. Review articles, commentaries, and appraisals were excluded. Only guidelines in English were included. The quality of the guidelines was assessed by one of the authors (WZ) using the AGREE instrument, which is specifically designed to assess the quality of guidelines.¹⁰ Quality score was calculated on a percentage scale for each domain as well as overall. Student's *t* test and one way analysis of variance were performed for two group and more than two group comparisons respectively. Mean (SD) or mean (95% confidence interval (95%CI)) were calculated as appropriate.

RESULTS

In addition to the EULAR recommendations, 1447 citations relating to hip and/or knee OA guidelines were retrieved from the literature search. After scrutiny, only 21 guidelines met the inclusion and exclusion criteria including the two EULAR recommendations.^{1-3, 6, 11-28}

Quality of the EULAR recommendations

Overall quality scores (maximum 100) of the EULAR recommendations were 51% and 57% for hip and knee OA respectively. They were in the highest quintile of the available guidelines (mean score 41%, range 9–65). The domain breakdown scores showed that the EULAR recommendations performed very well with respect to scope and purpose, rigour of development, and clarity. In contrast, both recommendations scored low for stakeholder involvement, applicability, and editorial independence (table 4).

Qualities of opinion based, evidence based, and hybrid guidelines

Guidelines may be categorised into three different types according to the source of evidence. If the source of evidence is predominantly derived from expert consensus, they are “opinion based guidelines”—for example, the Royal College of Physicians guidelines.²⁶ If the source of evidence is solely research evidence, they are termed “evidence based guidelines”—for example, the Prodigy guidelines.¹³ If both expert consensus and research evidence were used, the term “hybrid guidelines” is used—for example, the EULAR recommendations.¹⁻³ Overall, with the AGREE instrument, hybrid guidelines had the highest quality scores, followed by the evidence based guidelines and then the opinion based guidelines ($p < 0.0001$) (fig 1).

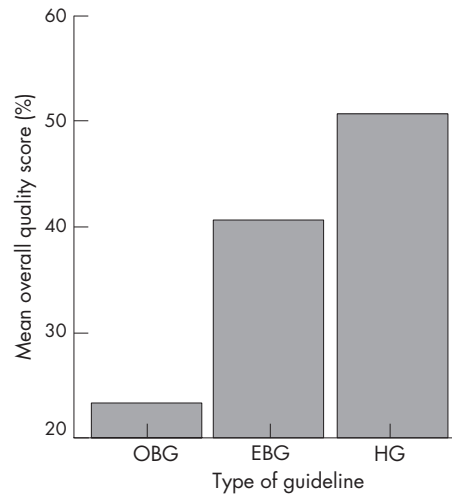


Figure 1 Overall quality score and type of guidelines. OBG, Opinion based guidelines; EBG, evidence based guidelines; HG, hybrid guidelines.

Strength of recommendation

Like many guidelines, the EULAR recommendations for the management of knee OA derived the strength of recommendation (SOR) according to the traditional method: the category of research evidence.¹ However, the EULAR Task Force altered this system when they developed guidelines for hip OA because they recognised inherent problems in this system. For example, for diacerhein and intra-articular steroid injection, randomised controlled trial (RCT) evidence was available but not supportive of these interventions for hip OA. Use of the traditional system would have resulted in a high SOR (grade A) for these treatments even though they are considered ineffective. In contrast, total hip replacement is recognised to be a clinically excellent treatment for severe hip OA, but for methodological and ethical reasons it has not been subjected to assessment by an RCT. Therefore total hip replacement could only be assigned a low SOR (grade C) even though it carried the full support of the Task Force.

Such caveats to the traditional SOR scale led the group to develop two alternative trade off scales: a visual analogue scale (VAS 0–100 mm) and an ordinal scale (A–E) for the SOR. Instead of assigning a SOR solely against the level of research evidence (in most cases, for efficacy only), the Task Force members were asked to mark their SORs on a 0–100 mm horizontal line with only two descriptive ends: “not recommended at all” and “fully recommended” taking into account the research evidence (for efficacy, safety, and cost effectiveness), clinical experience, logistical issues—for example, availability, cost, ease of delivery—and perceived patient acceptability and preference. They were also asked to select a discrete SOR from A (fully recommended), B

Table 4 Quality of the European League Against Rheumatism (EULAR) recommendations in the context of all guidelines in the management of hip and knee osteoarthritis (OA)

	Hip OA	Knee OA	Mean (range) of all OA guidelines
Scope and purpose	94.44	88.89	68.06 (22.22–100)
Stakeholder involvement	25.00	38.89	30.33 (0.00–62.50)
Rigour	71.43	80.95	39.77 (0.00–80.95)
Clarity	79.17	80.50	59.53 (31.25–91.67)
Applicability	22.22	37.04	18.06 (0.00–44.44)
Editorial independence	16.67	16.67	35.92 (0.00–88.89)
Overall	51.49	57.16	41.95 (9.26–65.34)

Values are AGREE scores (%).

(strongly recommended), C (moderately recommended), D (weakly recommended), and E (not recommended) in the same way. To date, two groups have used these trade off scales.³⁻⁶ The quality of the guidelines with these novel trade off scales tends to be higher numerically (fig 2).

The criterion validity of the trade off VAS has been examined.²⁹ As there is no gold standard in this area, the traditional scale was used as a proxy measure. The examination was undertaken within the same group of treatment modalities—exercise therapy for the management of hip and knee OA—where research evidence of different kinds is available and therefore both scales could be fairly compared. A significant linearity was observed between the trade off VAS and the traditional SOR ($p < 0.001$) (fig 3).²⁹

The reliability of the VAS and the ordinal scale has been examined using a test-retest method. Assessment of the SOR was repeated after two weeks by the same group of experts. The intraclass correlation coefficient was 0.60 (95%CI 0.52 to 0.67) for the VAS, and the weighted κ was 0.41 (95%CI 0.32 to 0.49) for the ordinal scale. The full details of the methods and results will be reported elsewhere.

The Delphi exercise

The EULAR recommendations used a Delphi exercise to reach expert consensus. This requires each Task Force member to propose, independently and away from a committee setting, an agreed number of key propositions for management—for

example, 10. The propositions from all members are then compiled into one list; at this stage an independent expert can amalgamate propositions that are very similar or overlap and can edit them for English and clarity (important for international committees). This first round list is then returned to all the Task Force members and they are asked to select their top favourites (a pre-defined number such as their top 10). Propositions are accepted if over half of the participants select them, whereas propositions receiving only a very few votes (a pre-defined number) are removed. Propositions receiving less than 50% but more than the minimum number enter the next Delphi round and members again vote for their top favourites. The procedure is repeated until the pre-defined number of propositions have been accepted. There are possible variations on this Delphi technique, but the key principles are:

- lack of influence from dominant individuals in an open committee setting
- acceptance of propositions by a majority decision
- equal weighting of all members with respect to proposing and voting.

When the key propositions have been agreed, the research evidence to support or refute each proposition is then examined using a systematic search strategy. The Task Force then discusses the evidence and determines the SORs.

Among 21 guidelines for OA, 15 used a consensus method: four of these used a Delphi exercise,^{1-3, 8, 20} two used conference consensus,²⁴⁻²⁷ and nine used methods that were not clearly defined.^{12, 15, 16, 19, 21, 23, 25, 26, 28} Interestingly the quality score of the guidelines using the Delphi technique appears to be greater than those with other consensus approaches (fig 4).

Other issues

Quality scoring of individual studies included in the research evidence, using a validated checklist, was undertaken by the EULAR OA Task Force in the development of their recommendations for management of knee OA.¹ However, they abandoned this practice when they developed recommendations for management of hip OA³ for the following reasons.

- “Quality scores” often judge the quality of reporting rather than the quality of the study design and the robustness of the evidence produced

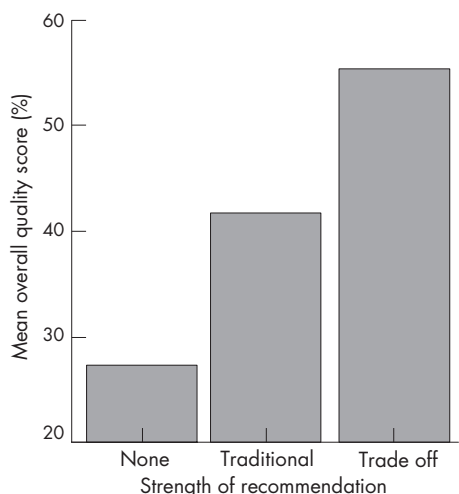


Figure 2 Overall quality score and strength of recommendation.

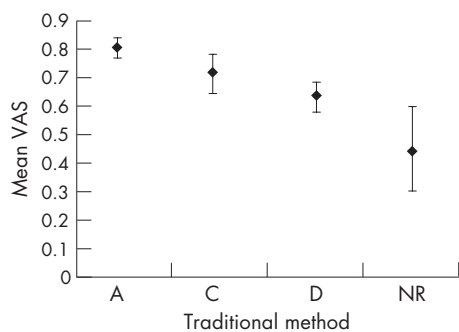


Figure 3 Relation between the traditional scale and the trade off visual analogue scale (VAS) for strength of recommendation. Data are mean and 95% confidence interval. A, level I evidence; B, level II or extrapolated from level I evidence; C, level III or extrapolated from level I or II evidence; D, level IV or extrapolated from level II or III evidence; NR, not recommended. Adapted with permission.²⁹

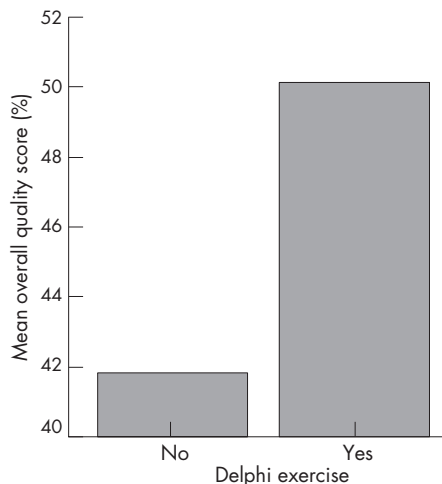


Figure 4 Overall quality score and Delphi exercise.

- With recent adoption of the CONSORT agreement (involving a checklist of key information for clinical trials), such quality scores have increased with recent publications and older publications tend to score much lower, making the quality score largely a function of date of publication
- Unless the quality scores are used to weight or influence in some way the category of evidence or SOR, it seems a relatively pointless, time consuming exercise; the panel members can judge the overall quality and robustness of the evidence without such a numerical score before deciding their SOR.

Three other guidelines^{6 15 19} assessed the quality of studies, but only one reported the quality scores for the evidence applied.⁶ Nevertheless, the quality of the guidelines was not affected by the quality scoring research evidence either for the overall score or individual domain scores of the AGREE instrument ($p > 0.05$).

The EULAR recommendations were developed by clinical experts—for example, rheumatologists, orthopaedic surgeons—with a special interest in OA. Therefore they are strongly biased towards a secondary, or even tertiary, care perspective, even though most of their recommendations are also applicable to primary care. No allied health professionals, general practitioners, or pharmacists were included in the Task Force even though such professionals are involved in care of patients with OA. Importantly, patient perspectives were also entirely unrepresented. The recommendations produced by EULAR are specific for hip or knee OA, not OA in general. In addition, the EULAR recommendations only highlight 10 key management issues; they are not intended as comprehensive guidelines for the complete management of patients with hip or knee OA (table 2).

DISCUSSION

The methodology involved in the development of treatment guidelines in OA has evolved rapidly during the last decade. Since the first OA guidelines were published by the Royal College of Physicians in 1993²⁶ through to 2005 when the EULAR recommendations were developed,³ the paradigm has shifted from opinion based to evidence based guidelines. The latter attempt to provide the best available evidence to support clinical decision making.³⁰ As a result, a number of evidence based guidelines have been developed for the management of hip and/or knee OA. Whereas some are solely evidence based such as the Prodigy guidance,¹³ others try to integrate both expert opinion and research evidence—that is, hybrid guidelines such as the EULAR recommendations.^{1 3} The overall quality of the guidelines has improved considerably (fig 1). This is mainly due to the improved scope and purpose, rigour of development, and editorial independence of the guidelines.

Interestingly, evidence based guidelines tend to have the lowest applicability, although the differences are not statistically significant (mean score: opinion based guidelines 17, evidence based guidelines 13, and hybrid guidelines 22, $p = 0.33$). This perhaps in part reflects the current difficulty in implementing evidence based guidelines, but also the gap between RCTs that demonstrate “efficacy” (an intervention works) and clinical “effectiveness” (how often and well the intervention will work in clinical practice). Although often regarded as the gold standard, RCTs have the following common problems.

- A focus on a highly selected, homogeneous sample of patients with knee or hip OA; often this means that the findings cannot be generalised to the whole population of patients with OA and clinical predictors of outcome cannot be examined

- Examination of a monotherapy rather than combined treatments or a package of care; interactions between treatments are under-studied
- Short study duration (a few weeks or months) even though symptomatic OA is usually a chronic condition.

Hybrid guidelines would be expected to improve applicability, as expert opinion can temper the rigidity of research data and close the gap between research and clinical practice.

The EULAR recommendations have claimed three strengths. Firstly, they align with the evidence based, clinical decision making scenario, a group decision based on research evidence, clinical expertise, and patient preference.³⁰ Although no patients were involved, their perceived preferences with respect to tolerability and acceptance were considered in the SOR. Secondly, EULAR recommendations use two established methods—systematic literature review and the Delphi exercise—to gather evidence from different sources—that is, research and clinical practice. Systematic review (or meta-analysis) is a well known and widely used quantitative method to synthesise research evidence, whereas the Delphi exercise is an established qualitative method to synthesise expert opinion.³¹ Both methods aim to obtain the best available evidence with the minimum individual (single study and single expert) bias. EULAR recommendations successfully apply both methods and thus substantially increase the quality of the guidelines, especially the rigour of the development (table 4). Finally, EULAR recommendations are supported by a trade off scale for SOR. Unlike the traditional SOR, which directly reflects the level of research evidence for efficacy,^{32 33} the trade off scale allows the developers to consider evidence from different sources (research and clinical) and to weigh, for example, benefits against harms and cost against effectiveness. The scales have been preliminarily validated with good criterion validity (fig 3) and moderate reliability.

However, EULAR recommendations also have shortcomings. Firstly, the guideline development team was biased by its predominance of rheumatologists and orthopaedic surgeons resulting in a lower score for stakeholder involvement (table 4). Although the recommendations have been adapted by many primary care groups, different recommendations may have been emphasised if general practitioners and allied health practitioners had been included in the Task Force. How to include patient representation in an international group representing many diverse countries is problematic,

What is already known on this topic

- EULAR has developed separate recommendations for management of hip and knee OA
- There are differences in detail from other guidelines on OA which largely result from the different evidence based formats used

What this study adds

- EULAR recommendations have high methodological rigour
- However, they are selective, not fully comprehensive, and both stakeholder involvement and editorial independence merit further improvement

but might best be achieved by a multi-country questionnaire survey rather than by representation of just a few select patients on the Task Force. Secondly, EULAR recommendations are not comprehensive guidelines for the management of knee or hip OA; only 10 key clinical questions have been addressed (table 2). Many other interventions were not considered, which may decrease the applicability of the recommendations. Thirdly, the meeting costs for the development of the EULAR recommendations were met by the pharmaceutical industry, and consequently the recommendations have lower editorial independence scores (16.67) (table 4).

This critical appraisal was based on the AGREE instrument. Although it is validated,³⁴ the AGREE instrument was developed at a time when expert opinion predominated guideline development, and research evidence was largely neglected. During our assessment, we experienced difficulty in assigning a score to some of the items and felt that apparently objective descriptions often required subjective judgment. Therefore we feel that some of the items in this instrument do not fairly reflect the true quality of guidelines and that further development of such an instrument is required. In addition, the guidelines included in this critical appraisal were assessed by only one person; because of the subjectivity of the instrument, the results have yet to be stabilised. Finally, the assessment only included OA guidelines in English. The quality of non-English guidelines remains unknown.

Overall, the EULAR recommendations have pioneered the development of hybrid guidelines, a possible future direction of clinical practice guidelines. Although limitations remain, the continuing evolution of the EULAR methodology provides an opportunity to improve the overall quality of guidelines for OA and other conditions.

Authors' affiliations

W Zhang, M Doherty, Academic Rheumatology, University of Nottingham, Nottingham City Hospital, Nottingham, UK

Competing interests: none declared

REFERENCES

- Jordan KM, Arden N, Doherty M, et al. EULAR recommendations 2003: an evidence based approach to the management of knee osteoarthritis: report of a Task Force of the Standing Committee for International Clinical Studies Including Therapeutic Trials (ESCSIT). *Ann Rheum Dis* 2003;**62**:1145–55.
- Pendleton A, Arden N, Dougados M, et al. EULAR recommendations for the management of knee osteoarthritis: report of a task force of the Standing Committee for International Clinical Studies Including Therapeutic Trials (ESCSIT). *Ann Rheum Dis* 2000;**59**:936–44.
- Zhang W, Doherty M, Arden N, et al. EULAR evidence based recommendations for the management of hip osteoarthritis: report of a task force of the EULAR Standing Committee for International Clinical Studies Including Therapeutics (ESCSIT). *Ann Rheum Dis* 2005;**64**:669–81.
- Mazieres B, Schmidely N, Hauselmann HJ, et al. Level of acceptability of EULAR recommendations for the management of knee osteoarthritis by practitioners in different European countries. *Ann Rheum Dis* 2005;**64**:1158–64.
- Punzi L, Canesi B, Carrabba M, et al. Italian consensus on Eular 2003 recommendations for the treatment of knee osteoarthritis. *Rheumatismo* 2004;**56**:190–201.
- Roddy E, Zhang W, Doherty M, et al. Evidence-based recommendations for the role of exercise in the management of osteoarthritis of the hip or knee: the MOVE consensus. *Rheumatology (Oxford)* 2005;**44**:67–73.
- Zochling J, van der Heijde D, Burgos-Vargas R, et al. ASAS/EULAR recommendations for the management of ankylosing spondylitis. *Ann Rheum Dis* 2006;**65**:442–52.
- Roddy E, Doherty M. Guidelines for management of osteoarthritis published by the American College of Rheumatology and the European League Against Rheumatism: why are they so different? *Rheum Dis Clin North Am* 2003;**29**:717–31.
- Svensson O, Malmenas M, Fajutrao L, et al. Greater reduction of knee than hip pain in osteoarthritis treated with naproxen, as evaluated by WOMAC and SF-36. *Ann Rheum Dis* 2006;**65**:781–4.
- The AGREE Collaboration. Appraisal of Guidelines for Research & Evaluation (AGREE) Instrument. <http://www.agreecollaboration.org> (accessed 9 Jun 2006).
- Algorithms for the diagnosis and management of musculoskeletal complaints. *Am J Med* 1997;**103**(suppl 1):S49–80.
- Universe of adult patients with osteoarthritis of the knee: Phase I, Phase II. American Academy of Orthopaedic Surgeons 2003. <http://www.aaos.org/> (accessed 9 Jun 2006).
- Prodigy guidance: osteoarthritis. Prodigy, The UK NHS 2005. <http://www.prodigy.nhs.uk/guidance.asp?gt=Osteoarthritis> (accessed 9 Jun 2006).
- Medical management of adults with osteoarthritis. Michigan Quality Improvement Consortium 2005. <http://www.mqic.org> (accessed 17 Oct 2005).
- Albright J, Allman R, Bonfiglio RP, et al. Philadelphia panel evidence-based clinical practice guidelines on selected rehabilitation interventions for knee pain. *Phys Ther* 2001;**81**:1675–700.
- Altman RD, Hochberg MC, Moskowitz RW, et al. Recommendations for the medical management of osteoarthritis of the hip and knee: 2000 update. *Arthritis Rheum* 2000;**43**:1905–15.
- Bennell K, Hinman R, Crossley K. APA knee joint osteoarthritis position statement. Australian Physiotherapy Association 2001. https://apa.advsol.com.au/staticcontent/staticpages/position_statements/mpa/kneeOASummary.pdf (accessed 9 Jun 2006).
- Brosseau L, Wells GA, Tugwell P, et al. Ottawa panel evidence-based clinical practice guidelines for therapeutic exercises and manual therapy in the management of osteoarthritis. *Phys Ther* 2005;**85**:907–71.
- 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**:526–30.
- Enloe LJ, Shields RK, Smith K, et al. Total hip and knee replacement programs: a report using consensus. *J Orthop Sports Phys Ther* 1996;**23**:3–11.
- Holbrook AM. Ontario treatment guidelines for osteoarthritis, rheumatoid arthritis and acute musculoskeletal injury. Ontario Program for Optimal Therapy 2000. <http://www.thecem.net/Downloads/msk.pdf> (accessed 9 Jun 2006).
- Lee J, Thorson D, Jurrison M, et al. Health care guideline: diagnosis and treatment of adult degenerative joint disease (DJD) of the knee. Institute for Clinical System Improvement. <http://www.icsi.org> (accessed 9 Jun 2006).
- Lundebjerg N. Exercise prescription for older adults with osteoarthritis pain: consensus practice recommendations. *J Am Geriatr Soc* 2001;**49**:808–23.
- Rankin EA, Alarcon GS, Chang RW, et al. NIH consensus statement on total knee replacement December 8–10, 2003. *J Bone Joint Surg [Am]* 2004;**86**:1328–35.
- Rosman HA, Tat KLK, et al. Clinical practice guidelines on the management of osteoarthritis. Academy of Medicine of Malaysia 2002. <http://www.acadmed.org.my/cpg/Bookleta.pdf> (accessed 9 Jun 2006).
- Scott DL, Billingham M, Bourke BE, et al. Guidelines for the diagnosis, investigation and management of osteoarthritis of the hip and knee: report of A Joint Working Group of the British Society for Rheumatology and the Research Unit of the Royal College of Physicians. *J R Coll Physicians Lond* 1993;**27**:391–6.
- Tannenbaum H, Peloso PM, Russell AS, et al. An evidence-based approach to prescribing NSAIDs in the treatment of osteoarthritis and rheumatoid arthritis: the Second Canadian Consensus Conference [review; 111 refs]. *Can J Clin Pharmacol* 2000;**7**(suppl A):4A–16A.
- Vogels EMHM, Hendriks HJM, van Barr ME, et al. Clinical practice guidelines for physical therapy in patients with osteoarthritis of the hip and knee. <http://www.cebp.nl/media/m11.pdf> (accessed 9 Jun 2006).
- Roddy E, Zhang W, Doherty M, et al. Evidence-based clinical guidelines: a new system to better determine true strength of recommendation. *J Eval Clin Pract* 2006;**12**:347–52.
- Hynes RB, Sackett DL, Gray JMA, et al. Transferring evidence from research into practice. 1. The role of clinical care research evidence in clinical decisions. *Evid Based Med* 1996;**1**:196–7.
- Jones J, Hunter D. Qualitative research: consensus methods for medical and health services research. *BMJ* 1995;**311**:376–380.
- The periodic health examination. Canadian Task Force on the Periodic Health Examination. *Can Med Assoc J* 1979;**121**:1193–254.
- Shekelle PG, Woolf SH, Eccles M, et al. Clinical guidelines: developing guidelines [review; 17 refs]. *BMJ* 1999;**318**:593–6.
- The AGREE Collaboration. Development and validation of an international appraisal instrument for assessing the quality of clinical practice guidelines: the AGREE project. *Qual Saf Health Care* 2003;**12**:18–23.