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 (ESCISIT)


Objectives: To develop evidence based recommendations for the management of hip osteoarthritis (OA).

Methods: The multidisciplinary guideline development group comprised 18 rheumatologists, 4 orthopaedic surgeons, and 1 epidemiologist, representing 14 European countries. Each participant contributed up to 10 propositions describing key clinical aspects of hip OA management. Ten final recommendations were agreed upon using a Delphi consensus approach. Medline, Embase, CINAHL, Cochrane Library, and HTA reports were searched systematically to obtain research evidence for each proposition. Where possible, outcome data for efficacy, adverse effects, and cost effectiveness were abstracted. Effect size, rate ratio, number needed to treat, and incremental cost effectiveness ratio were calculated. The quality of evidence was categorised according to the evidence hierarchy. The strength of recommendation was assessed using the traditional A–D grading scale and a visual analogue scale.

Results: Ten key treatment propositions were generated through three Delphi rounds. They included 21 interventions, such as paracetamol, NSAIDs, symptomatic slow acting disease modifying drugs, opioids, intra-articular steroids, non-pharmacological treatment, total hip replacement, osteotomy, and two general propositions. 461 studies were identified from the literature search for the proposed interventions of efficacy, adverse effects, and cost effectiveness. Research evidence supported 15 interventions in the treatment of hip OA. Evidence specific for the hip was strikingly lacking. Strength of recommendation varied according to category of research evidence and expert opinion.

Conclusion: Ten key recommendations for the treatment of hip OA were developed based on research evidence and expert consensus. The effectiveness and cost effectiveness of these recommendations were evaluated and the strength of recommendation was scored.

METHODS

Participants

A multidisciplinary guideline development committee was commissioned by ESCISIT. Twenty three experts in the field of OA (18 rheumatologists, 4 orthopaedic surgeons, and 1 epidemiologist) representing 14 European countries agreed to take part in the study. The objectives were (a) to agree 10 key propositions for the management of hip OA; (b) to identify and critically appraise research evidence for the effectiveness of interventions; (c) to evaluate the strength of evidence and to develop the recommendations.

Abbreviations: ASU, avacado soybean unsaponifiable; CI, confidence interval; COX-2, cyclo-oxygenase-2; CS, chondroitin sulphate; CT, controlled trial; CV, cardiovascular; ES, effect size; GI, gastrointestinal; GS, glucosamine sulphate; HA, hyaluronic acid; ICER, incremental cost effectiveness ratio; NNT, number needed to treat; NSAIDs, non-steroidal anti-inflammatory drugs; QALY, quality of life year; OA, osteoarthritis; OR, odds ratio; PPI, proton pump inhibitor; RCT, randomised controlled trial; RR, relative risk; SYSDOA, symptomatic slow acting drugs for OA; THR, total hip replacement; VAS, visual analogue scale.
and cost effectiveness of the relevant treatments; and (c) to
generate recommendations based on a combination of the
best available evidence and expert opinion.

Experts’ consensus
Each committee member was asked to contribute independ-
ently up to 10 propositions describing key clinical interven-
tions for OA of the hip. Consensus on the propositions was
reached using the Delphi technique. The initial propositions
were collated into a single list. Similar, overlapping proposi-
tions were combined. The list was returned to the experts and
they were asked to select the 10 most important from the list.
Propositions were accepted automatically if selected by over
half of participants in any round, whereas propositions
receiving only one to three votes were removed. Propositions
receiving less than 50% of the votes but more than three votes
entered the next Delphi round. The procedure was repeated
until 10 propositions were agreed.

Systematic literature search
A systematic search of the literature published between
January 1966 and March 2004 was undertaken using
Medline, Embase, CINHAL, and Cochrane Library databases.
The search included both a general search and an interven-
tion-specific search. The general search strategy consisted of
two basic components: hip OA in whatever possible terms
in the databases (Appendix 1, available at http://www.
anrnheumdis.com supplemental) and types of research in
the forms of systematic review/meta-analysis, randomised
controlled trial (RCT)/controlled trial (CT), uncontrolled trial,
cohort study, case-control study, cross sectional study,
and economic evaluation (Appendix 2, available at http://
www.anrnheumdis.com supplemental). The general search
aimed at summarising the current available reported treat-
ments for hip OA. The summary results of this search were
reported to the committee before the Delphi exercise.

After the Delphi exercise had generated the 10 proposi-
tions, the intervention-specific search was undertaken to
identify evidence for each specific intervention. The search
strategy included the terms for hip OA (Appendix 1, http://
www.anrnheumdis.com supplemental) and any possible
terms for the specific intervention. For example paracetamol,
acetaminophen, or simple analgesics were used for parac-
etamol. The results of the two searches were then combined
and duplications excluded. Medical subject heading search
(MeSH) was used for all databases and a keyword search was
used if the MeSH search was not available. All MeSH search
terms were exploded. The reference lists within reviews or
systematic reviews were examined and any additional studies
meeting the inclusion/exclusion criteria were included.

The search in the Cochrane Library included MeSH search
of the Cochrane Review, Abstracts of Quality Assessed
Systematic Reviews, the Cochrane Controlled Trial Register,
NHS Economic Evaluation Databases, the Health Technology
Assessment Database, and NHS Economic Evaluation Bibliography Details only. In addition, a topics search on
OA was undertaken.

Inclusion/exclusion criteria
Only studies with clinical outcomes for hip OA were
included. The main focus of interest was on systematic
reviews, RCTs/CTs, uncontrolled trials/cohort studies, case-
control studies, cross sectional studies, and economic
evaluations. Studies that combined hip and knee (and/or
other joint) OA were excluded unless the results for patients
with hip OA were reported separately. Studies of other sites
of OA or other chronic joint conditions were excluded unless
the adverse effects were investigated as a primary outcome.
Other exclusions were case reports, animal studies, non-
clinical outcome studies, narrative review articles, com-
mentaries, and guidelines.

Categorising evidence
Evidence was categorised according to study design using a
hierarchy of evidence in descending order according to
quality6 (table 1). Questions on efficacy were answered
using the best available evidence. For example, if a question
on the effect of an intervention could be answered by
category Ia evidence (that is, a systematic review of RCTs)
then studies of a weaker design (RCTs, category Ib) were not
reviewed. Questions on side effects were answered using both
RCTs and observational studies. Although the efficacy was
assessed specifically for hip OA, side effects were evaluated
specifically for the intervention irrespective of the musculo-
skeletal condition. Questions on cost effectiveness were
answered according to the outcome measure of the effec-
tiveness. For example, if the effectiveness was measured as
pain relief or quality of life years (QALYs) gained, only
studies for hip OA were eligible. If the effectiveness was
measured as adverse events averted, any studies for the
proposed interventions were included.

Table 1 Evidence hierarchy and traditional strength of recommendation

<table>
<thead>
<tr>
<th>Category of evidence</th>
<th>Strength of recommendation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ia Meta-analysis of RCTs</td>
<td>A Category I evidence</td>
</tr>
<tr>
<td>Ib RCT</td>
<td></td>
</tr>
<tr>
<td>IIB Controlled study without randomisation</td>
<td>B Category II evidence or extrapolated from category I evidence</td>
</tr>
<tr>
<td>IIB Quasi-experimental study</td>
<td></td>
</tr>
<tr>
<td>III Non-experimental descriptive studies, such as comparative, correlation, and case-control studies</td>
<td>C Category III evidence or extrapolated from category I or II evidence</td>
</tr>
<tr>
<td>IV Expert committee reports or opinion or clinical experience of respected authorities, or both</td>
<td>D Category IV evidence or extrapolated from category II or III evidence</td>
</tr>
</tbody>
</table>

Estimation of effectiveness and cost effectiveness
Effect size (ES) and 95% confidence interval (95% CI)
compared with placebo or active control as specified within
the propositions was calculated for continuous outcomes
such as pain relief and improvement in function.9 The ES is
the standard mean difference—that is, the mean difference
between a treatment and a control group, divided by the
standard deviation of the difference. It therefore has no units
and is comparable across the interventions. Clinically,
ES = 0.2 is considered small, ES = 0.5 is moderate, and
ES >0.8 is large.10 Results from the latest systematic review
were used if more than one systematic review was available
for the same intervention. Statistical pooling was undertaken
as appropriate11 when a systematic review was not available.
The percentage of patients with moderate to excellent (or >50%) pain relief or symptomatic improvement was calculated and the number needed to treat (NNT) was estimated. A positive value for the NNT means the treatment is more beneficial than control; whereas a negative value for the NNT means the treatment is less beneficial than control. NNT and 95% CI were reported only if it was statistically significant; otherwise ‘NS’ (not significant) was used to avoid the confusion due to its unique mathematical features. Relative risk (RR) was calculated for adverse effects.

When economic evaluation was available, study design, comparator, perspective, time horizon, discounting, total costs, effectiveness were reviewed. The incremental cost effectiveness ratio (ICER) was calculated. QALYs were used when available, otherwise disease-specific outcomes such as pain relief and functional improvement were used.

Data were extracted by two investigators using a customised form. Any discrepancies were discussed and agreed between the extractors before analysis. Non-English language studies were extracted by the native investigators.

Strength of recommendation
The strength of recommendation was graded A–D based on the category of efficacy evidence (table 1) by two members of the committee (WZ, MD) and subsequently ratified by the committee. When hip-specific data were absent and therefore no hip-specific category was applicable, no strength of recommendation could be applied. However, a visual analogue scale (VAS) was also used. Each member of the committee was asked to mark on a 0–100 mm VAS their strength of recommendation for each intervention, according to the research evidence (efficacy, safety, and cost effectiveness) and clinical expertise (logistics, patient perceived acceptance and tolerability). The means and standard error of the mean (SEM) for the strength of recommendation were calculated for each intervention. This system allowed strength of recommendation to be applied when hip-specific efficacy data were absent.

Future research agenda
Each committee member was asked to propose 10 topics for the future research agenda based on current available evidence and clinical experience in the management of hip OA. A Delphi approach was used to reach a consensus on the 10 most important topics. The same criteria as those used to select propositions were employed (that is, accepted when
Treatment modalities and types of research evidence

The general search yielded 1725 hits (Medline 1143, Embase 357, CINAHL 46, and Cochrane 179). After deleting duplications, 1341 hits remained. Of these, 898 were original studies and 443 were narrative reviews, commentary, or editorials. Figure 1 shows the breakdown of interventions among the original 898 studies and fig 2 shows the types of evidence.

Experts’ opinion approach

The experts were informed of the results of the general literature search and then the Delphi exercise was undertaken. One hundred and twelve propositions were produced initially and the 10 final propositions were agreed after three anonymous Delphi rounds (table 2).

Assessment of propositions

The results from the intervention-specific search were merged with the results from the general search. After deleting the duplications and articles irrelevant to the questions, 461 studies remained on the list. These included 44 concerning paracetamol; 287 associated with conventional non-steroidal anti-inflammatory drugs (NSAIDs), cyclo-oxygenase-2 (COX-2) selective inhibitors (coxibs), and gastroprotective agents; 41 in relation to symptomatic slow acting drugs for OA (SYSADOA); 26 concerning opioid analgesics; 7 for intra-articular steroid injection; 22 related to non-pharmacological treatments; and 34 relevant to the two surgical propositions. The following propositions are grouped by topic (general, non-pharmacological, pharmacological, invasive and surgical) with no weighting according to order.

1. The optimal management of hip OA requires a combination of non-pharmacological and pharmacological treatment modalities

Although this statement is logical and represents common clinical practice, there are no direct comparisons or evidence from appropriately designed clinical trials using a factorial design to support this statement. The statement is supported by expert opinion alone (category IV).

2. The treatment of hip OA should be tailored according to:
   - Hip risk factors (obesity, adverse mechanical factors, physical activity, dysplasia)
   - General risk factors (age, sex, comorbidity, co-medication)
   - Level of pain intensity, disability, and handicap
   - Location and degree of structural damage
   - Wishes and expectations of the patient

This statement represents ideal practice and includes clinical markers that are often used to guide clinical decisions. However, although it has considerable commonsense face validity, there is little experimental evidence to support it. RCTs predominantly investigate the efficacy of one or two

Table 2 Experts’ propositions developed through three Delphi rounds—order according to topic (general, non-pharmacological, pharmacological, invasive and surgical)

<table>
<thead>
<tr>
<th>No</th>
<th>Proposition</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Optimal management of hip OA requires a combination of non-pharmacological and pharmacological treatment modalities</td>
</tr>
<tr>
<td>2</td>
<td>Treatment of hip OA should be tailored according to: (a) Hip risk factors (obesity, adverse mechanical factors, physical activity, dysplasia) (b) General risk factors (age, sex, comorbidity, co-medication) (c) Level of pain intensity, disability, and handicap (d) Location and degree of structural damage (e) Wishes and expectations of the patient</td>
</tr>
<tr>
<td>3</td>
<td>Non-pharmacological treatment of hip OA should include regular education, exercise, appliances (stick, insoles), and weight reduction if obese or overweight</td>
</tr>
<tr>
<td>4</td>
<td>Because of its efficacy and safety paracetamol (up to 4 g/day) is the oral analgesic of first choice for mild-moderate pain and, if successful, is the preferred long term oral analgesic</td>
</tr>
<tr>
<td>5</td>
<td>NSAIDs, at the lowest effective dose, should be added or substituted in patients who respond inadequately to paracetamol. In patients with increased gastrointestinal risk, non-selective NSAIDs plus a gastroprotective agent, or a selective COX-2 inhibitor (coxib) should be used</td>
</tr>
<tr>
<td>6</td>
<td>Opioid analgesics, with or without paracetamol, are useful alternatives in patients in whom NSAIDs, including COX-2 selective inhibitors (coxibs), are contraindicated, ineffective, and/or poorly tolerated</td>
</tr>
<tr>
<td>7</td>
<td>SYSADOA (glucosamine sulphate, chondroitin sulphate, diacerein, avocado soybean unsaponifiable, and hyaluronic acid) have a symptomatic effect and low toxicity, but effect sizes are small, suitable patients are not well defined, and clinically relevant structure modification and pharmacoeconomic aspects are not well established</td>
</tr>
<tr>
<td>8</td>
<td>intra-articular steroid injections (guided by ultrasound or x-ray) may be considered in patients with a flare that is unresponsive to analgesic and NSAIDs</td>
</tr>
<tr>
<td>9</td>
<td>Osteotomy and joint preserving surgical procedures should be considered in young adults with symptomatic hip OA, especially in the presence of dysplasia or varus/valgus deformity</td>
</tr>
<tr>
<td>10</td>
<td>Joint replacement has to be considered in patients with radiographic evidence of hip OA who have refractory pain and disability</td>
</tr>
</tbody>
</table>

OA, osteoarthritis; NSAIDs, non-steroidal anti-inflammatory drugs; COX-2, cyclo-oxygenase-2; SYSADOA, symptomatic slow acting drugs for osteoarthritis.
specific monotherapies in highly selected homogeneous populations of otherwise fit subjects with hip OA. The evidence obtained from such experimental studies therefore may not be directly applicable to the whole population of subjects with hip OA. In addition, because of exclusion of many variables that may influence outcome it is often difficult to determine predictors of outcome. Therefore, better evidence may be obtained from observational studies, such as the large number of cohort studies undertaken to observe possible risk factors or predictors at baseline for total hip replacement (THR) or osteotomy. For example, age, level of dysplasia, and degree of deformity are major predictors for THR III 118 2–20 years – – –
Osteotomy III 9 2–20 years – – –
Intra-articular steroid Ib 1 3 months – – –
Hyaluronic acid III 3 3–12 months – – –
Avocado soybean unsaponifiable Ib 2 6–12 months 0.31 (0.23 to 0.47) 0.31 (0.23 to 0.47) –
Diclofenac Ib 1 3 years 0.00 (–0.17 to 0.17) 0.00 (–0.17 to 0.17) –
Intra-articular steroid Ib 1 3 months – – –
Osteotomy III 3 2–20 years – – –
THR III 118 2–20 years – – –

*See Table 1 for definitions.
No, number of studies; ES, effect size of treatment compared with placebo unless otherwise stated; NNT, number needed to treat to obtain moderate to excellent (≥50%) pain relief or symptomatic improvement; –, not available; NS, not significant.

In summary, only one RCT (category Ib) has been undertaken for education in hip OA alone and it suggests that education reduces pain. However, evidence (category Ia) for OA at any joint showed that education may have very little value and may not reach statistical significance. In contrast, although direct evidence for hip OA is lacking, exercise appears to be beneficial for OA of any kind (category Ia). There is some evidence to support the benefit of weight reduction for hip OA (category III), but no evidence for appliances (stick and insoles) (table 3). Nevertheless, despite the absence of trial data, interventions that reduce adverse mechanical forces across a compromised hip joint have obvious face validity. 4. Because of its efficacy and safety paracetamol (up to 4 g/day) is the oral analgesic of first choice for mild-moderate pain and, if successful, is the preferred long term oral analgesic. There have been no placebo controlled RCTs, and no comparative efficacy studies in hip OA alone, although paracetamol is widely prescribed for hip OA. Three systematic reviews were identified. The most recent covered four placebo controlled trials, of which two were undertaken in patients with hip or knee OA over a 6 week period. The effect size for pain relief was 0.21 (95% CI 0.02 to 0.41) and the NNT to obtain clinical benefit, defined as moderate to excellent pain relief from paracetamol over placebo, was 4 (95% CI 2 to 43). However, paracetamol was inferior to, but safer than, NSAIDs.

Table 3 Evidence of efficacy—pooled effect size (ES) and number needed to treat (NNT)
Recently, concern has been raised over the possible GI toxicity of paracetamol. One case-control study indicated that paracetamol taken at a dose of more than 2 g daily was associated with a greater risk of GI perforation or bleed (OR = 3.6, 95% CI 2.6 to 5.1), and one cohort study reported a dose-response relationship between paracetamol and dyspepsia. However, evidence from a systematic review of RCTs shows that paracetamol has no more GI upsets than placebo (RR = 0.80, 95% CI 0.27 to 2.37). A meta-analysis of case-control studies also showed no increased risk of GI bleeding with paracetamol 2–4 g daily (OR = 1.2, 95% CI 0.8 to 1.7) and no dose dependent effect in the range of <2 g, 2–4 g, and >4 g per day, an observation that fits with endoscopic studies.

Although there are some concerns about renal toxicity of long term regular use of paracetamol, evidence to support this is sparse and conflicting (table 4). In conclusion, there is no direct evidence to support the use of paracetamol in hip OA. However, evidence in OA of any site (category Ia) demonstrates that paracetamol is effective in relieving pain arising from OA. Although it is inferior to conventional NSAIDs, it is safer when taken within the recommended dose range. There are no long term data on the efficacy or safety of paracetamol and no direct evidence for its cost effectiveness in the treatment of hip OA. However, a study in knee OA shows that paracetamol is better than NSAIDs, NSAIDs plus gastroprotective agents, and coxibs for the cost of each GI adverse effect avoided (table 5).

### Table 4 Evidence of safety—pooled relative risk (RR*) and 95% confidence interval (CI)

<table>
<thead>
<tr>
<th>Intervention</th>
<th>Adverse events</th>
<th>RR (95% CI)</th>
<th>Category of evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Paracetamol</td>
<td>GI discomfort</td>
<td>0.80 (0.27 to 2.37)</td>
<td>RCTs</td>
</tr>
<tr>
<td></td>
<td>GI perforation/bleed</td>
<td>3.60 (2.60 to 5.10)</td>
<td>Case-control study</td>
</tr>
<tr>
<td></td>
<td>GI bleeding</td>
<td>1.2 (0.8 to 1.7)</td>
<td>Case-control studies</td>
</tr>
<tr>
<td></td>
<td>Renal failure</td>
<td>0.83 (0.50 to 1.39)</td>
<td>Cohort study</td>
</tr>
<tr>
<td></td>
<td>Renal failure</td>
<td>2.5 (1.7 to 3.6)</td>
<td>Case-control study</td>
</tr>
<tr>
<td>NSAIDs</td>
<td>GI perforation/ulcer/bleed</td>
<td>5.36 (1.79 to 16.10)</td>
<td>RCTs</td>
</tr>
<tr>
<td></td>
<td>GI perforation/ulcer/bleed</td>
<td>2.70 (2.10 to 3.50)</td>
<td>Cohort studies</td>
</tr>
<tr>
<td></td>
<td>GI perforation/ulcer/bleed</td>
<td>3.00 (2.70 to 3.70)</td>
<td>Case-control studies</td>
</tr>
<tr>
<td>Coxibs vs NSAIDs</td>
<td>Endoscopic GI ulcer</td>
<td>0.18 (0.14 to 0.23)</td>
<td>RCTs</td>
</tr>
<tr>
<td></td>
<td>CV events</td>
<td>0.79 (0.40 to 1.55)</td>
<td>RCTs</td>
</tr>
<tr>
<td>Coxibs vs naproxen</td>
<td>CV events</td>
<td>1.69 (1.07 to 2.69)</td>
<td>RCTs</td>
</tr>
<tr>
<td>Misoprostol</td>
<td>Endoscopic GI ulcer</td>
<td>0.26 (0.17 to 0.39)</td>
<td>RCTs</td>
</tr>
<tr>
<td></td>
<td>Diarrhoea</td>
<td>1.81 (1.52 to 2.16)</td>
<td>RCTs</td>
</tr>
<tr>
<td>H2 Blockers (double doses)</td>
<td>Endoscopic GI ulcer</td>
<td>0.44 (0.03 to 0.74)</td>
<td>RCTs</td>
</tr>
<tr>
<td>PPIs</td>
<td>Endoscopic GI ulcer</td>
<td>0.40 (0.32 to 0.51)</td>
<td>RCTs</td>
</tr>
<tr>
<td>Opioids+paracetamol vs paracetamol</td>
<td>GI upset/congestion</td>
<td>14.00 (1.86 to 105.16)</td>
<td>RCT</td>
</tr>
<tr>
<td>Dicacehin</td>
<td>Diarrhoea</td>
<td>3.73 (2.61 to 5.32)</td>
<td>RCT</td>
</tr>
<tr>
<td></td>
<td>Skin rash/pruritus</td>
<td>2.40 (1.01 to 5.69)</td>
<td>RCT</td>
</tr>
</tbody>
</table>

*RR, relative risk between treatment group and control group; RR = 1, no different from the control population; RR > 1, more risky than the control population; RR < 1, less risky than the control population. Pooled RR was estimated for more than one study. Compared with placebo/non-exposure unless otherwise stated.

### Table 5 Evidence of cost effectiveness for the proposed interventions

<table>
<thead>
<tr>
<th>Intervention</th>
<th>Comparator</th>
<th>Perspective</th>
<th>Duration</th>
<th>Discounting</th>
<th>Effectiveness</th>
<th>C1-C2</th>
<th>E1-E2</th>
<th>ICER</th>
</tr>
</thead>
<tbody>
<tr>
<td>Paracetamol</td>
<td>Ibuprofen</td>
<td>Institutional/payer</td>
<td>6 months</td>
<td>No</td>
<td>PUB overtaken</td>
<td>$63000–112000</td>
<td>995–980</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>Rofecoxib</td>
<td>Institutional/payer</td>
<td>6 months</td>
<td>No</td>
<td>PUB overtaken</td>
<td>$63000–471000</td>
<td>995–991</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>Celecoxib</td>
<td>Institutional/payer</td>
<td>6 months</td>
<td>No</td>
<td>PUB overtaken</td>
<td>$63000–474000</td>
<td>995–990</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>Misoprostol</td>
<td>Institutional/payer</td>
<td>6 months</td>
<td>No</td>
<td>PUB overtaken</td>
<td>$471000–112000</td>
<td>991–980</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>NSAIDs</td>
<td>Institutional/payer</td>
<td>1 year</td>
<td>3%</td>
<td>PUB overtaken</td>
<td>–</td>
<td>–</td>
<td>4738</td>
</tr>
<tr>
<td></td>
<td>Dicacehin</td>
<td>Institutional/payer</td>
<td>6 months</td>
<td>No</td>
<td>PUB overtaken</td>
<td>$474000–112000</td>
<td>990–980</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>NSAIDs</td>
<td>Institutional/payer</td>
<td>6 months</td>
<td>No</td>
<td>PUB overtaken</td>
<td>$556000–112000</td>
<td>988–980</td>
<td>54146</td>
</tr>
<tr>
<td></td>
<td>NSAIDs</td>
<td>Canadian Health services</td>
<td>3 months</td>
<td>No</td>
<td>PUB overtaken</td>
<td>$32396–25622</td>
<td>96–86</td>
<td>684</td>
</tr>
<tr>
<td>NSAIDs+misoprostol (all patients with OA)</td>
<td>NSAIDs</td>
<td>Institutional/payer</td>
<td>3 years</td>
<td>No</td>
<td>PUB overtaken</td>
<td>$556000–112000</td>
<td>988–980</td>
<td>54146</td>
</tr>
<tr>
<td>NSAIDs+misoprostol (patients with OA aged &gt;65 years)</td>
<td>NSAIDs</td>
<td>Institutional/payer</td>
<td>3 years</td>
<td>No</td>
<td>PUB overtaken</td>
<td>$556000–112000</td>
<td>988–980</td>
<td>54146</td>
</tr>
<tr>
<td>Dicacehin+standard treatment</td>
<td>NSAIDs</td>
<td>Institutional/payer</td>
<td>9 months</td>
<td>No</td>
<td>Lequesne function</td>
<td>$2360–2272</td>
<td>2.2–2.12</td>
<td>88</td>
</tr>
<tr>
<td>THR (women, 60 years)</td>
<td>Conventional treatment</td>
<td>Societal</td>
<td>Life</td>
<td>5%</td>
<td>QALYs</td>
<td>$47649–165440</td>
<td>4.16–2.16</td>
<td>17121</td>
</tr>
<tr>
<td>THR (men, &gt;85 years)</td>
<td>Conventional treatment</td>
<td>Societal</td>
<td>Life</td>
<td>5%</td>
<td>QALYs</td>
<td>$30580–21432</td>
<td>4.16–2.16</td>
<td>4754</td>
</tr>
<tr>
<td>THR No THR</td>
<td>Societal</td>
<td>1 year</td>
<td>No</td>
<td>QALYs</td>
<td>$9990–0</td>
<td>0.84–0.29</td>
<td>18164</td>
<td></td>
</tr>
<tr>
<td>THR No THR</td>
<td>NHS</td>
<td>Life</td>
<td>6%</td>
<td>QALYs</td>
<td>$4804–0</td>
<td>8.39–0</td>
<td>573</td>
<td></td>
</tr>
<tr>
<td>THR No THR</td>
<td>Irish institutional</td>
<td>2 years</td>
<td>No</td>
<td>QALYs</td>
<td>$572.06–0</td>
<td>74.2–48.8</td>
<td>19</td>
<td></td>
</tr>
</tbody>
</table>

C1, total costs with intervention; C2, total costs with comparator; E1, effect with intervention; E2, effect with comparator; ICER, incremental cost effectiveness ratio; base case scenario; PUB, perforation, ulcer, or bleed; QALYs, quality adjusted life years.
5. NSAIDs, at the lowest effective dose, should be added or substituted in patients who respond inadequately to paracetamol. In patients with increased gastrointestinal risk, non-selective NSAIDs plus a gastroprotective agent, or a selective COX-2 inhibitor (coxib) should be used

One systematic review has been undertaken for non-aspirin NSAIDs in the treatment of hip OA, in which 14 placebo controlled trials were reviewed.47 The ES for pain relief was 0.69 (95% CI 0.12 to 1.26) and the NNT to obtain clinical benefit over placebo was 4 (95% CI 3 to 6), supporting the use of NSAIDs for hip OA (table 3). However, the GI safety of NSAIDs counteracts their benefits. Six systematic reviews have been undertaken to assess the GI safety of NSAIDs using evidence from RCTs, cohort, and case-control studies.48–52 The results show that NSAIDs cause an increased risk of GI bleeding (table 4), which is dose dependent (for example, OR = 2.2 (95% CI 0.8 to 5.8), 3.2 (1.9 to 5.5), and 12.2 (5.6 to 26.7) for diclofenac <75 mg, 75–150 mg, and >150 mg a day, respectively).48

A number of strategies have been used to minimise the GI risk due to NSAIDs. Two systematic reviews demonstrated a significant reduction of GI toxicity with coxibs,53–55 but five systematic reviews provided evidence to support the co-administration of non-selective conventional NSAIDs with gastroprotective agents such as misoprostol, double doses of H2 blockers, and proton pump inhibitors (PPIs).56–60 Table 4 presents the pooled relative risks of endoscopic gastric ulcer between coxibs or co-administration of gastroprotective agents and NSAIDs from the latest systematic review.56 A similar pattern is seen for other definitions of peptic ulcer and ulcer complications, but standard dose H2 blockers are less effective than other GI protectors, and care must be taken when using misoprostol as it causes an increase in diarrhoea (RR = 1.81 95% CI 1.52 to 2.16).56

In addition, there is concern over potential cardiovascular (CV) side effects (for example, myocardial infarction or stroke) of rofecoxib. However, a systematic review of 23 clinical trials showed that rofecoxib is only associated with a greater CV risk in comparison with naproxen (RR = 1.69, 95% CI 1.07 to 2.69). The risk was not statistically significant when compared with placebo (RR = 0.84 95% CI 0.51 to 1.38), or non-naproxen NSAIDs (RR = 0.79, 95% CI 0.40 to 1.53).56 The extra CV thrombotic events with rofecoxib may therefore reflect more the antiplatelet or some other protective effects of naproxen than a side effect of rofecoxib. Further investigation of this issue is continuing.

The other important issue related to the use of coxibs or GI protectors is whether the extra benefits can balance the extra cost. Three economic evaluations of the treatment of OA at any site have been undertaken.61–64 Because they all use clinical GI events (perforation, ulcer, or bleed) averted as an effectiveness outcome measure, the results may be generalisable to patients with hip OA. The results from one study showed that the ICER of coxibs versus ibuprofen was $31 769.9 for each adverse event averted for rofecoxib and $33 518.5 for celecoxib. The ICER reduced when the patients’ GI risk increased, indicating that the use of coxibs is more cost effective for patients with a higher risk of GI bleeding. Similar results were obtained with GI protective agents but with a greater ICER, suggesting that co-prescription of GI protectors is the more expensive strategy (table 5).

In summary, NSAIDs are effective in relieving the pain of hip OA (category Ia). However, the GI side effects of NSAIDs offset their benefits (category Ia). Although coxibs or the addition of GI protectors (misoprostol, double doses of H2 blockers, and PPIs) to conventional NSAIDs can significantly reduce GI bleeding (category Ia), these strategies are more expensive and only cost effective in patients with greater GI risk.

6. Opioid analgesics, with or without paracetamol, are useful alternatives in patients in whom NSAIDs, including COX-2 selective inhibitors (coxibs) are contraindicated, ineffective, and/or poorly tolerated

One systematic review reported that a single dose of a combination of paracetamol and codeine increases by about 5% the analgesic strength of treatment of any type of pain, including pain due to hip OA.65 Unfortunately, the review failed to separate hip OA from other conditions. The review additionally showed that paracetamol plus codeine caused more adverse events than paracetamol alone (RR = 2.5, 95% CI 1.5 to 4.2). Four RCTs were identified for hip OA, one comparing codeine with placebo66 and three comparing paracetamol plus opioid with placebo,67 paracetamol,68 or diclofenac.68 Opioid (codeine) on its own was better than placebo for pain relief (ES = 0.78, 95% CI 0.38 to 1.18) and functional improvement (ES = 0.78, 95% CI 0.26 to 1.30), but it also caused more adverse events (RR = 1.43, 95% CI 1.09 to 1.86).67 Although the combination of paracetamol and opioid provided better analgesia than placebo (ES = 0.30, 95% CI 0.05 to 0.53),69 this treatment was no better than paracetamol (RR for moderate to excellent pain relief 1.27, 95% CI 0.82 to 1.98)67 and was inferior to diclofenac (ES = −0.18, 95% CI −0.33 to −0.04).68 More importantly, the combination caused more side effects (RR = 7.25, 95% CI 2.61 to 20.13), GI upsets, constipation (RR = 14, 95% CI 1.86 to 105.16), and dizziness or drowsiness (RR = 5.00, 95% CI 1.48 to 16.92)70 and resulted in greater withdrawal rates (RR = 3.57, 95% CI 1.92 to 6.62)67 (table 4).

In summary, opioid analgesics with or without paracetamol are effective for hip OA (category Ib). However, the effect may be no better than paracetamol alone and inferior to NSAIDs (category Ib). In addition, they cause more side effects and resultant cessation of treatment (category Ia and Ib). Such evidence supports the statement and reinforces the suggestion that opioid analgesics should only be considered for patients in whom paracetamol, conventional NSAIDs or coxibs have insufficient efficacy, or in whom there are contraindications to NSAIDs/coxibs.

7. SYSADOA (glucosamine sulphate, chondroitin sulphate, diacerein, avocado soybean unsaponifiable, and hyaluronic acid) have a symptomatic effect and low toxicity, but effect sizes are small, suitable patients are not well defined, and clinically relevant structure modification and pharmacoeconomic aspects are not well established

The term SYSADOA covers a range of agents, including glucosamine sulphate (GS), chondroitin sulphate (CS), diacerein, avocado soybean unsaponifiable (ASU), and hyaluronic acid (HA). The classification varies from country to country. In the UK, for example, GS and CS are classified as health food supplements and are available over the counter. A number of systematic reviews have been undertaken to support the use of oral GS71–74 and CS.75–77 However, none of them are specific for the hip. In the mixed trials with hip and knee OA, the pooled ES of GS versus placebo were 0.44 (95% CI 0.24 to 0.64) for pain relief and 0.41 (95% CI 0.14 to 0.69) for functional improvement, while those for CS were 0.78 (95% CI 0.60 to 0.95) and 0.63 (95% CI 0.32 to 0.94), respectively. One hip-specific placebo controlled RCT was undertaken for CS, demonstrating that CS was statistically better than placebo in reducing pain and improving function over 6 months of treatment. However, the ES could not be calculated as standard deviations were not reported.70 The systematic literature search failed to identify other levels of evidence specific for hip OA. Evidence for structure modifying effects of these two agents has not been
established. One RCT demonstrated that intramuscular glycosaminoglycan peptide complex (now withdrawn from the market) given twice a year for 5 years had no more structural benefits than placebo for hip OA.73

Two systematic reviews72 73 have been undertaken for herbal treatment, including two RCTs74 75 for ASU in OA of various sites. A subgroup of patients with hip OA (n = 50) were available from one of the trials with a 24 week treatment period.72 The results showed a significantly greater pain relief than placebo (ES = 0.76, 95% CI 0.17 to 1.34). However, this was not supported by a recent trial with a larger sample size (n = 163) and longer treatment period (1 year for symptomatic outcomes and 2 years for structural changes).76 The pooled ES for pain relief and functional improvement of these two trials were not statistically significant (table 3). In addition, there were no structural benefits with ASU over the 2 year treatment period.76

Evidence for diacerhein is inconclusive. One multicentre RCT with 507 patients with hip OA showed that diacerhein had no more pain relief (ES = 0.00, 95% CI −0.17 to 0.17) or functional improvement (ES = 0.00, 95% CI −0.17 to 0.17) than placebo over a 3 year treatment period.77 However, this study was designed to investigate structure modification rather than symptom benefit. Another RCT of 207 patients with hip or knee OA demonstrated that the combination of diacerhein and standard treatment was more effective in reducing pain (ES = 0.29, 95% CI 0.05 to 0.57) and improving function (ES = 0.35, 95% CI 0.08 to 0.63) than standard treatment alone over a 6 month period,78 though data for hip OA alone was not presented.

However, regardless of efficacy diacerhein causes an increase in the incidence of side effects such as diarrhoea (RR = 3.73, 95% CI 2.61 to 5.32), and skin rash or pruritus (RR = 2.40, 95% CI 1.01 to 5.69).77 Nevertheless, diacerhein may slow the progression of joint space narrowing in hip OA.77 The relative risk of the progression, defined as a joint space narrowing ≥0.5 mm, during a 3 year treatment period compared with placebo was 0.84 (95% CI 0.71 to 0.99), with an NNT of 10 (95% CI 5 to 171). However, this was not supported by a recent trial with a larger sample size (n = 163) and longer treatment period (1 year for symptomatic outcomes and 2 years for structural changes).76 The pooled ES for pain relief and functional improvement of these two trials were not statistically significant (table 3). In addition, there were no structural benefits with ASU over the 2 year treatment period.76

In conclusion, there is no direct evidence to support the effectiveness of SYSADOA have yet to be established. One RCT (category Ib) demonstrates that CS effectively reduces pain and functional disability due to hip OA. The symptomatic benefits of ASU and diacerhein are inconclusive (category Ib) and the evidence for HA is poor (category III). The structure modifying effect and cost effectiveness of SYSADOA have yet to be established.

8. Intra-articular steroid injections (guided by ultrasound or X ray) may be considered in patients with a flare that is unresponsive to analgesic and NSAIDs

Three trials have been undertaken for hip OA,82–84 but only one is an RCT comparing a steroid-anaesthetic combination with anaesthetic alone.85 Only a dichotomous outcome is available from this trial and therefore the ES was not calculated. The rate ratio for pain relief was 1.18 (95% CI 0.68 to 2.15) and 0.61 (0.23 to 1.60) at 1 month and 3 months, respectively, indicating that the combination was no better than anaesthetic alone for hip OA. Although two uncontrolled trials showed some short term (≤3 months) pain reduction from intra-articular corticosteroid,83 84 the results are open to bias due to placebo effects. Thus unlike for knee OA there is no robust evidence to support the efficacy of steroid injection for hip OA, irrespective of its increased technical difficulty. Also, there are no comparative data to show increased accuracy using ultrasound or X ray guidance, and no data available for the occurrence of a flare.

In conclusion, there is category Ib evidence for this statement, but the results are inconclusive and placebo controlled trials in hip OA are required.

9. Osteotomy and joint preserving surgical procedures should be considered in young adults with symptomatic hip OA, especially in the presence of dysplasia or varus/valgus deformity

Osteotomies of the pelvis and/or femoral osteotomies can alter force transmission through the hip joint and thus potentially influence clinical symptoms and the course of the OA process. Symptomatic and structural effects have been evaluated, especially in patients with acetabular dysplasia and valgus deformity of the femoral neck. As with THR, however, most published investigations are cohort studies, and we failed to identify any RCT for either symptoms or structural change. At least seven retrospective cohort studies85 86 and two prospective cohort studies87 88 with follow up times ranging from 2 to 20 years have investigated both outcomes and baseline predictors relevant to the proposition. Patients showed significant improvement in

| Table 6 | Radiographic severity, pain, and function and relative risk of total hip replacement |
|-----------------|-----------------|------------------|-----------------|
| Risk factors    | Cohort studies  | Case-control studies | |
|                 | No of studies   | RR (95% CI)       | No of studies   | OR (95% CI) |
| Overall x ray change* | 4 | 2.39 (1.74 to 3.29) | 1 | 1.98 (1.23 to 3.19) |
| Croft grade     |           |                  |                |
| 0/1             | 1          | 1                |                |
| 2               |            | 3.36 (0.31 to 38.91) |               |
| 3               |            | 15.23 (3.29 to 70.49) |               |
| 4               |            | 44.51 (10.04 to 197.48) |               |
| 5               |            | 57.29 (12.12 to 270.71) |               |
| Pain [≥50%]     | 1          | 1.86 (1.23 to 3.88) | 1 | 1.91 (1.43 to 2.56) |
| Lequesne function [≥10] | 2 | 2.75 (1.98 to 3.82) | 1 | 1.75 (1.24 to 2.48) |

*Defined as Croft grade ≥2, or joint space of <2 mm. All values are adjusted by age, sex, and body mass index. A pooled RR/OR was estimated for more than one study.

www.annrheumdis.com
clinical outcomes (pain, walking ability, and overall functional scores) and radiographic outcome after the operation. Advanced age,\textsuperscript{15 91} radiographic severity,\textsuperscript{16 85–88} degree of dysplasia,\textsuperscript{15 88} and radiographic deformity\textsuperscript{15 16} at baseline were associated with worse clinical outcomes and with surgical failure, mainly defined as requirement for THR. As the vast majority of studies have been performed in young adults with at least mild or moderate symptoms (mainly hip pain) and data on the natural course of femoral and acetabular deformities is lacking, the effectiveness of osteotomy in asymptomatic patients and in different age groups has yet to be established.

Data from observational studies also support more recently advocated joint preserving surgical procedures such as arthroscopic debridement\textsuperscript{92 93} and surgical dislocation of the hip with offset reconstruction.\textsuperscript{94 95} Although improvement of symptoms has been reported in these studies, the lack of control groups with an alternative treatment hinders interpretation of their results.

In conclusion, evidence for osteotomy and joint preserving surgical procedures in patients with hip OA is sparse (category III). It appears to be a useful procedure for younger patients with painful hip dysplasia or deformity for whom THR is not yet justified. However, its effectiveness and cost effectiveness as compared with THR in patients with advanced age and/or OA stages have yet to be established.

10. Joint replacement has to be considered in patients with radiographic evidence of hip OA who have refractory pain and disability

Because of methodological and ethical problems, comparison of THR with placebo or standard care is not readily evaluated by an RCT. Our literature search failed to identify any placebo or standard care controlled RCT for THR. However, there are a large number of head to head comparisons between different types of prosthesis and uncontrolled follow up studies.\textsuperscript{96 97} One systematic review identified 118 uncontrolled follow up studies involving a total of 77,375 patients for an average follow up period of 9.4 years (range 2–20 years).\textsuperscript{97} The percentage of patients free from pain at the end point ranged from 43.2% (95% CI 34 to 49) to 84.1% (95% CI 46 to 100) depending upon the type of prosthesis. The mean reduction of Harris hip score (maximum 100, including pain 0–40, function 0–47, motion 0–5, and deformity 0–8) from baseline ranged from 36% to 46%. The revision rate, one of the major concerns for THR, was 0.18 (standard error of mean (SEM) 0.04) to 2.04 (SEM 0.19) per 100 person-years adjusted by age, sex, and type of hip arthritis. More information about the difference between prostheses is available,\textsuperscript{96 97} but this falls beyond the scope of this statement.

For concurrent comparison, better pain relief and quality of life were obtained with THR in one small Asian cohort study of THR versus non-surgical treatment. However, although in the first year THR was better than non-surgical treatment, by 5 years this benefit was reversed.\textsuperscript{98} Because the study had a very small sample size (62 v 45) and no adjustment for confounders, the results are subject to bias. Two other cohort studies have been undertaken, one comparing WOMAC and SF-36 scores\textsuperscript{99} and the other comparing survival rates\textsuperscript{100} between patients with THR and the source populations. Apart from worse functional scores, the THR group had a similar quality of life and survival rate to those of the source populations.

Whether radiographic change should be one of the criteria for THR is still open to debate. Six cohort studies\textsuperscript{101–106} (four with data available for re-analysis) and one case-control study\textsuperscript{107} demonstrated a positive relationship between radiographic severity and risk of THR (table 6). A greater risk of THR was also seen with severe pain and disability. Thus the severity of radiographic change, pain, and disability are good predictors for THR. However, whether they are independent risk factors for THR or are factors that surgeons currently use to guide the decision process is hard to determine. Nevertheless, as long as the diagnosis of OA is confirmed (most commonly by radiographic change) it has obvious validity for the severity of patient centred problems (pain, disability) to largely determine the need for major surgery.

Another important consideration for THR is whether it is cost effective compared with other treatments. Five economic evaluations of THR were identified, though one compared different prostheses\textsuperscript{108} so is not relevant to this statement. Four studies compared THR either with conventional

<p>| Table 7 | Future research agenda—propositions developed through three Delphi rounds |</p>
<table>
<thead>
<tr>
<th>No</th>
<th>Proposition</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>More RCTs of both pharmacological and non-pharmacological treatments that give outcomes specific to hip OA are needed</td>
</tr>
<tr>
<td>2</td>
<td>Biological markers for evaluation of the progression of hip OA should be further evaluated</td>
</tr>
<tr>
<td>3</td>
<td>Clinical predictors of response to pharmaceutical and non-pharmacological interventions for hip OA should be determined</td>
</tr>
<tr>
<td>4</td>
<td>Whether long term use of SYSADOA can retard the progression of hip OA and delay joint replacement should be investigated</td>
</tr>
<tr>
<td>5</td>
<td>RCTs of injection treatments (corticosteroid, hyaluronan) in hip OA are required</td>
</tr>
<tr>
<td>6</td>
<td>The most efficient and effective exercise programme for hip OA should be determined</td>
</tr>
<tr>
<td>7</td>
<td>Studies with appropriate design to determine the comparative effectiveness and cost effectiveness of non-surgical and surgical treatment modalities are needed</td>
</tr>
<tr>
<td>8</td>
<td>Agreed criteria relating to indications and timing of THR are needed</td>
</tr>
<tr>
<td>9</td>
<td>Prospective population based studies are required to improve our knowledge of risk factors for the development and progression of hip OA</td>
</tr>
<tr>
<td>10</td>
<td>Newer imaging techniques (MRI, ultrasound) require validation for the diagnosis and assessment of outcome in trials of hip OA</td>
</tr>
</tbody>
</table>
and they are directly associated with postoperative appearance the most important and agreed measures for THR. While the importance of the degree of disability to be key determinants of the need for major surgery (category III). In conclusion, THR is effective in hip OA in improving clinical outcomes such as pain and function (category III). In general, THR is more cost effective in younger women. Radiographic change is usually sufficient to confirm the diagnosis of hip OA, and the severity of radiographic change, as well as pain and disability are risk factors for THR. There is obvious face validity and evidence for the severity of pain and disability to be key determinants of the need for major surgery (category III).

### Future research agenda

Ninety five research topics were recommended initially. Table 7 shows the 10 that were agreed eventually as the most important topics for future research according to current available research evidence and clinical practice.

#### DISCUSSION

This is the first comprehensive document to provide recommendations for the management of hip OA. Unlike previous guidelines that are specific for or inclusive of hip OA, these recommendations are based both on expert opinion and research evidence, with clear separation between the two. We have employed explicit methods such as the Delphi technique to generate consensus and an evidence based medicine approach to identify and appraise the research evidence. A similar hybrid technique was used to develop the EULAR recommendations for the management of knee OA. However, for these recommendations on hip OA we introduced three important methodological changes.

Firstly, we did not score the quality of the studies. We had found this exercise to be unhelpful in assessing the research evidence because quality scores are subject to bias that results from the quality of reporting. For example, RCTs that appear from RCTs and applied the VAS 0–100 mm scale for the strength of recommendation to facilitate the multidimensional issues for each treatment or non-THR treatment, assuming there was no cost or effects for the non-THR group. The results showed that THR did require additional cost to gain its additional benefits over conventional treatment or non-THR treatment (table 5). However, it was more cost effective in younger women. For example, compared with conventional treatment THR was cost saving for a woman aged 60 years (ICER $17121 per QALY gained), whereas it was less cost effective for a man aged more than 85 years (ICER $4,754 per QALY gained).

Given the benefits and costs of THR, determination of the ideal point at which to perform surgery in the course of OA is crucial. A number of priority criteria have been proposed for patients awaiting THR, but none of them are universally agreed. While the importance of the degree of radiographic change remains unclear, pain, and function appear the most important and agreed measures for THR and they are directly associated with postoperative outcomes.

In conclusion, THR is effective in hip OA in improving clinical outcomes such as pain and function (category III). In general, THR is more cost effective in younger women. Radiographic change is usually sufficient to confirm the diagnosis of hip OA, and the severity of radiographic change, as well as pain and disability are risk factors for THR. There is obvious face validity and evidence for the severity of pain and disability to be key determinants of the need for major surgery (category III).

#### Table 8 Strength of recommendation

<table>
<thead>
<tr>
<th>Intervention</th>
<th>Research evidence*</th>
<th>Side effects</th>
<th>Cost effectiveness</th>
<th>SOR based on efficacy (A–D)</th>
<th>SOR based on all evidence and clinical expertise (VAS, mean (SEM))</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pharmacological + non-pharmacological treatment</td>
<td>IV +</td>
<td>–</td>
<td>–</td>
<td>D</td>
<td>86.94 (5.82)</td>
</tr>
<tr>
<td>Treatment tailored according to risk factors, severity of hip OA, and patient expectations Education</td>
<td>III +</td>
<td>–</td>
<td>–</td>
<td>C</td>
<td>92.19 (3.39)</td>
</tr>
<tr>
<td>Exercise</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>A</td>
<td>71.75 (6.42)</td>
</tr>
<tr>
<td>Insole/stick</td>
<td>IV +</td>
<td>–</td>
<td>–</td>
<td>D</td>
<td>61.72 (6.91)</td>
</tr>
<tr>
<td>Weight loss</td>
<td>III +</td>
<td>–</td>
<td>–</td>
<td>D</td>
<td>68.28 (5.79)</td>
</tr>
<tr>
<td>Paracetamol</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>D</td>
<td>79.19 (3.82)</td>
</tr>
<tr>
<td>NSAIDs</td>
<td>la +</td>
<td>la +, III +</td>
<td>Cost saving</td>
<td>N/A</td>
<td>79.36 (4.18)</td>
</tr>
<tr>
<td>Coxibs</td>
<td>la + (GI protection)</td>
<td>la ± (CV)</td>
<td>Higher GI risk population</td>
<td>A</td>
<td>79.44 (3.51)</td>
</tr>
<tr>
<td>Misoprostol</td>
<td>la + (GI protection)</td>
<td>la + (diarrhoea)</td>
<td>Higher GI risk population</td>
<td>A</td>
<td>46.06 (5.62)</td>
</tr>
<tr>
<td>H2 Blockers (double dose)</td>
<td>la + (GI protection)</td>
<td>–</td>
<td>High GI risk population</td>
<td>A</td>
<td>31.28 (6.81)</td>
</tr>
<tr>
<td>PPIs</td>
<td>la + (GI protection)</td>
<td>–</td>
<td>–</td>
<td>A</td>
<td>73.86 (3.97)</td>
</tr>
<tr>
<td>Opioids</td>
<td>la + (any, GI, CNS)</td>
<td>–</td>
<td>–</td>
<td>A</td>
<td>43.97 (4.36)</td>
</tr>
<tr>
<td>Glucosamine</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>N/A</td>
<td>37.06 (5.03)</td>
</tr>
<tr>
<td>Chondroitin</td>
<td>ib +</td>
<td>–</td>
<td>–</td>
<td>A</td>
<td>34.44 (4.76)</td>
</tr>
<tr>
<td>Diacerein</td>
<td>ib ±</td>
<td>ib + (dyspepsia)</td>
<td>Short term Inconclusive</td>
<td>27.83 (5.38)</td>
<td></td>
</tr>
<tr>
<td>Avocado soybean unsaponifiable</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>Not recommended</td>
<td>31.72 (4.79)</td>
</tr>
<tr>
<td>Hyaluronic acid</td>
<td>III +</td>
<td>–</td>
<td>–</td>
<td>C</td>
<td>22.83 (4.17)</td>
</tr>
<tr>
<td>Intra-articular steroid</td>
<td>ib –</td>
<td>–</td>
<td>–</td>
<td>Not recommended</td>
<td>41.47 (5.74)</td>
</tr>
<tr>
<td>Osteotomy</td>
<td>III +</td>
<td>–</td>
<td>–</td>
<td>C</td>
<td>59.64 (5.19)</td>
</tr>
<tr>
<td>THR</td>
<td>III +</td>
<td>–</td>
<td>–</td>
<td>Women with younger age</td>
<td>86.86 (2.42)</td>
</tr>
</tbody>
</table>

*Evidence was categorised according to the hierarchy in table 1. **-** not supportive, **+** supportive, **±** uncertain. For example, la + (GI) means there is category Ia evidence to support the statement that the treatment causes GI side effects.

SOR, strength of recommendation; VAS, visual analogue scale (0–100 mm, 0 = not recommended at all, 100 = fully recommended); SEM, standard error; GI, gastrointestinal; CV, cardiovascular; CNS, central nervous system; –, not available; N/A, not applicable owing to absent hip-specific data.
Of the 21 interventions included within the 10 propositions, 15 were positively supported by evidence of grades Ia to IV (table 8), but 6 of them had either no direct evidence (paracetamol, glucosamine, and exercise) or inconclusive benefits (ASU, diacerein, and intra-articular steroid injection) for hip OA. In contrast, these treatments are effective and have been recommended for knee OA. More data on these interventions for hip OA are required, but the current evidence suggests that there may be true treatment differences for OA according to the site affected. Such differences support the requirement by regulatory bodies to obtain separate research evidence for the benefit of treatments at each key site of OA. Similar site-specific differences for OA and hip OA may occur for risk factors for development and progression of OA and for the correlation between pain and structural change, reflecting the heterogeneity of the OA process.

Although the evidence hierarchy is widely used to rank the quality of evidence, its value in surgical treatment and side effects has been questioned. THR, for example, has been accepted as a clinically effective treatment for hip OA, particularly for the patient with refractory pain and disability, or for those who fail to respond to conventional treatment. However, because of ethical and practical issues over blinding of the treatments, there are no placebo controlled or concurrent non-surgical controlled RCTs. The evidence to support THR comes from uncontrolled or cohort studies and is thus graded as category III, which is discordant with the high strength of support for THR when all forms of evidence, not just research evidence, are considered (table 8). Furthermore, in the absence of hip-specific data and therefore no category of evidence, strength of recommendation based on the evidence hierarchy could not be applied. Such problems challenge the current closely linked methods for categorising evidence and affording a strength of recommendation.

To overcome this, we undertook a post hoc study by asking the committee members to mark the strength of recommendation for each intervention on a VAS (0–100 mm). Compared with the traditional grading scale for strength of recommendation, the VAS scale considers research evidence of all kinds (efficacy, safety, and cost effectiveness) and clinical expertise. More importantly, the VAS scale allows both downgrading and upgrading of the strength of recommendation, offering a different dimension from the category of evidence. For example, the strength of recommendation for NSAIDs is 79% given the highest category of evidence (Ia) for efficacy, whereas the strength of recommendation for total hip replacement is 86%, although it only obtained category III evidence for efficacy (table 8).

These recommendations have several limitations. Firstly, although management of hip OA was the primary interest, for some interventions, such as paracetamol and exercise, we failed to identify any studies specific to hip OA or studies in which the data for hip OA could be separated. In this situation, we specified lack of hip-specific data but considered the evidence from mixed studies when determining the strength of recommendation. This may cause some imprecision and the values are yet to be confirmed. Secondly, the effect sizes were selected from the latest systematic review with the maximum number of the studies involved, but not necessarily the review with the best quality and relevancy. Subgroup analyses were often required, but in most cases the data were not available. Thirdly, only effect sizes for symptomatic outcomes such as pain and function were examined for efficacy. Efficacy beyond these patient centred outcomes remains unknown. Finally, the existing evidence hierarchy centres on treatment efficacy, whereas evidence for safety and cost effectiveness is best examined by designs other than RCTs and therefore requires its own specific system to judge category of evidence. In the absence of such a specific system we used the existing hierarchy, but would favour future development of an equivalent system for these other aspects of management.

In conclusion, we have developed recommendations for the management of hip OA based on both clinical practice and the best available research evidence. Fifteen of 21 interventions included within 10 key propositions have research evidence (category Ia–IV) to support their use in the management of hip OA, although they vary in their efficacy, side effect, and cost effectiveness profiles. Three interventions (paracetamol, GS, and exercise) had no hip-specific evidence and three interventions (ASU, diacerein, and intra-articular injection) had category I evidence to show either no symptomatic benefit or inconclusive evidence for hip OA, even though these treatments are effective for knee OA. Clearly, more clinical trial data specific to hip OA are required, especially because some interventions appear to show different efficacy according to the joint site. It is hoped that wide dissemination and discussion of these recommendations within healthcare provider groups will improve knowledge and interest in the management of hip OA and result in higher standards of care for patients with hip OA.

ACKNOWLEDGEMENTS
We thank Bristol Myers Squibb, in particular its representative—Dr Manuela Le Bars, for financial support, and Mrs Helen Richardson and Dr Jinying Lin for logistical support.

Authors’ affiliations
M Dougados, Service de Rhumatologie B, Hospital Cochin, Paris, France
M Doherty, W Zhang, Academic Rheumatology, University of Nottingham, Nottingham, UK
N Arden, K Jordan, Southampton General Hospital, MRC Environmental Epidemiology Unit, Southampton, UK
B Bannwarth, Service de Rhumatologie, Hôpital Pellegrin, Bordeaux, France
J Bijlsma, Department of Rheumatology and Immunology, University Hospital, Utrecht, The Netherlands
K-P Gunther, Department of Orthopaedic Surgery, University of Dresden, Dresden, Germany
H J Haussmann, Centre for Rheumatology and Bone Diseases, Clinic Im Park, University of Zurich, Zurich, Switzerland
G Herrero-Beaumont, Rheumatology Department, Clinique de la Conception, Madrid, Spain
P Kaklamani, 16 Anaperon Polemon, 11521, Athens, Greece
B Leeb, Rheumatology, Stockeyu Hospital, Stockeuy, Austria
M Lequesne, Department of Rheumatology, Hôpital Léopold Bellan, 75014, Paris, France
S Lohmander, Department of Orthopaedics, Lund University Hospital, SE-22185, Lund, Sweden
B Mazieres, Service de Rhumatologie, Hôpital de Rangueil, Toulouse, France
E Martin-Mola, Division de Rhumatologia, Hospital La Paz, Madrid, Spain
P Pawlak, Institute of Rheumatology, 12850, Praha 2, Czech Republic
A Pendleton, Rheumatology, Belfast City Hospital, Belfast, UK
L Punzi, Department of Medical Sciences, University of Padova, Padova, Italy
B Swoboda, Orthopaedics, University of Erlangen-Nuremberg, Erlangen, Germany
R Varação, No32B 1400–107, Lisbon, Portugal
G Verbruggen, Rheumatology Unit, UZBRUG, Gent, Belgium
I Zimmermann-Gorska, Department of Rheumatology and Rehabilitation and Internal Medicine, Poznań, University of Medical Sciences, Czerna, Poland

REFERENCES


