EXTENDED REPORT

Quality indicators for the primary care of osteoarthritis: a systematic review

J J Edwards,1 M Khanna,2 K P Jordan,1 J L Jordan,1 J Bedson,1 K S Dziedzic1

ABSTRACT

Objective To identify valid and feasible quality indicators for the primary care of osteoarthritis (OA).

Design Systematic review and narrative synthesis.

Data sources Electronic reference databases (MEDLINE, EMBASE, CINAHL, HMIC, PsychINFO), quality indicator repositories, subject experts.

Eligibility criteria Eligible articles referred to adults with OA, focused on development or implementation of quality indicators, and relevant to UK primary care. An English language restriction was used. The date range for the search was January 2000 to August 2013. The majority of OA management guidance has been published within this time frame.

Data extraction Relevant studies were quality assessed using previous quality indicator methodology. Two reviewers independently extracted data. Articles were assessed through the Outcome Measures in Rheumatology filter; indicators were mapped to management guidance for OA in adults. A narrative synthesis was used to combine the indicators within themes.

Results 10 853 articles were identified from the search; 32 were included in the review. Fifteen indicators were considered valid and feasible for implementation in primary care; these related to assessment non-pharmacological and pharmacological management. Another 10 indicators were considered less feasible, in various aspects of assessment and management. A small number of recommendations had no published corresponding quality indicator, such as use of topical non-steroidal anti-inflammatory drugs. No negative (‘do not do’) indicators were identified.

Conclusions and implications of key findings There are well-developed, feasible indicators of quality of care for OA which could be implemented in primary care. Their use would assist the audit and quality improvement for this common and frequently disabling condition.

BACKGROUND

Osteoarthritis (OA) is a common reason for consultation with a general practitioner (GP): around 4% of the population aged 45 years and over will consult a GP in a year with a diagnosis of OA.1 One working definition of OA is “persistent joint pain that is worse with use [in people] age 45 years old and over [who have] morning stiffness lasting no more than half an hour” and does not require radiography for diagnosis.2 There are evidence-based interventions to reduce pain and disability in adults with OA. Guidance on the care and management of OA has been produced by the American College of Rheumatology, the European League Against Rheumatism, the Osteoarthritis Research Society International, and the National Institute for Health and Care Excellence (NICE).3–8 Although management may vary by the site of OA, core aspects of primary care management are generally common across all sites.4–6 8 If these interventions were routinely implemented by GPs, there would be a significant impact on population levels of pain and disability attributable to OA.9 However, there is evidence that such implementation is not occurring.10–15

Routine audit and feedback on provided care is needed to improve the quality of that care. Quality indicators (hereafter ‘indicators’) are one suitable tool.16 Such indicators are defined as a “measurable [element] of practice performance for which there is evidence or consensus that it can be used to assess the quality, and hence change in the quality, of care provided”.17 Although reviews by Hochberg18 and Strömbeck et al19 identified indicators for measuring quality of care for OA, which show promise for use in primary care, there has been no systematic review and synthesis of the development and implementation literature to identify the most promising and feasible set of primary care OA indicators. Hunter et al20 argue cogently for ‘further systematic development, implementation, and audit of quality measures.’ The objective of this systematic review was to identify existing indicators of core treatment for OA feasible for use in primary care medical records and for routine audit purposes through electronic data retrieval.

METHODS

We used the methodology for systematic reviews set out by the Centre for Reviews and Dissemination.21

Review protocol

Available on request from the corresponding author.

SEARCH STRATEGY

A search strategy was developed to identify articles concerning the development, testing or implementation of indicators of the quality of care for OA applicable to adults in a primary medical care setting.

The systematic search strategy was customised for use in databases searchable through the UK National Health Service (NHS) Evidence portal (CINAHL, EMBASE, HMIC, MEDLINE and PsychINFO). A range of OA terms were combined with indicator terms. An English language restriction was used. The date range for the search was 2000 to 2013.

Special search for non-English literature

The search strategy was customised for non-English language restriction was used. The date range for the search was 2000 to 2013.
The 10 groups of studies in which indicators for OA care had been developed are listed in table 2. The following aspects of quality assessment were common to all studies and are not included in the table.

Although not every study explicitly declared there to be no conflict of interest, the reviewers considered that no significant resulting bias of the results was likely.

No studies had an identified method of updating the indicators in light of new evidence.

External validity and sensitivity to change had not been demonstrated in any of the indicator development, testing or implementation studies.

Reproducibility, at the level of the individual indicator, is shown in table 3.

Of 10 indicator development study groups, five were based on the Assessing Care of Vulnerable Elders (ACOVE) indicators. Overall, the ACOVE series of indicators were found to most closely fulfil the assessment criteria due to their robust evidence collection and consensus development, and field testing, and update in ACOVE-3. The modifications to ACOVE-1 indicators (for use the English Longitudinal Study of Aging,41 in nursing homes42 and home-based primary care45) were minor, such as to the target population or recommended care process time frames. The degree to which modifications were subject to further empirical study and consensus varied. We judged the modified indicators to be compatible with the originals, although there was variability regarding the indicators of use of oral non-steroidal anti-inflammatories (NSAIDs) and gastroprotective agents, in terms of the drugs recommended or the target population. The RAND indicators were the earliest identified; they were based on a literature review (not identified as systematic) and high-quality consensus exercise. The developers of the Arthritis Foundation indicators had undertaken a ‘comprehensive’ literature review, and a high-quality consensus exercise. One example of implementation (of the non-pharmacological indicators) was found. The remaining indicator sets used an evidence synthesis or consensus exercise which was less rigorous, or not specified. Some had no identified evidence of implementation (eg, Physician Consortium for Performance Improvement (PCPI) indicators).

All identified articles used process-of-care measures as indicators; one indicator set (European Musculoskeletal Conditions Surveillance and Information Network; EUMUSC.net) also used three outcome measures. We identified no papers in which quality improvement over time had been investigated.

### Results

#### Selection of articles

Ten thousand eight hundred and fifty-two unique articles were identified. The final inclusion set numbered 32. There were 10 groups of indicators in 14 development articles, and 18 implementation articles.

The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flow chart setting out the review process can be found in the online supplementary figure S2. Excluded studies are listed in the online supplementary table S3.

#### Assessment of quality

The 10 groups of studies in which indicators for OA care had been developed are listed in table 2. The following aspects of quality assessment were common to all studies and are not included in the table.

Although not every study explicitly declared there to be no conflict of interest, the reviewers considered that no significant resulting bias of the results was likely.

No studies had an identified method of updating the indicators in light of new evidence.

External validity and sensitivity to change had not been demonstrated in any of the indicator development, testing or implementation studies.

Reproducibility, at the level of the individual indicator, is shown in table 3.

Of 10 indicator development study groups, five were based on the Assessing Care of Vulnerable Elders (ACOVE) indicators. Overall, the ACOVE series of indicators were found to most closely fulfil the assessment criteria due to their robust evidence collection and consensus development, and field testing, and update in ACOVE-3. The modifications to ACOVE-1 indicators (for use the English Longitudinal Study of Aging,41 in nursing homes42 and home-based primary care45) were minor, such as to the target population or recommended care process time frames. The degree to which modifications were subject to further empirical study and consensus varied. We judged the modified indicators to be compatible with the originals, although there was variability regarding the indicators of use of oral non-steroidal anti-inflammatories (NSAIDs) and gastroprotective agents, in terms of the drugs recommended or the target population. The RAND indicators were the earliest identified; they were based on a literature review (not identified as systematic) and high-quality consensus exercise. The developers of the Arthritis Foundation indicators had undertaken a ‘comprehensive’ literature review, and a high-quality consensus exercise. One example of implementation (of the non-pharmacological indicators) was found. The remaining indicator sets used an evidence synthesis or consensus exercise which was less rigorous, or not specified. Some had no identified evidence of implementation (eg, Physician Consortium for Performance Improvement (PCPI) indicators).

All identified articles used process-of-care measures as indicators; one indicator set (European Musculoskeletal Conditions Surveillance and Information Network; EUMUSC.net) also used three outcome measures. We identified no papers in which quality improvement over time had been investigated.

---

**Table 1** MEDLINE search strategy

| Number | MEDLINE expression | Search terms |
|--------|---------------------|--------------|
| 1      | ((qualit* ADJ3 (outcome* OR indicat*))).ti,ab | 1 MEDLINE | 15 OR 16 |
| 2      | QUALITY OF HEALTH CARE/ | 2 MEDLINE | osteoarth* .ti,ab |
| 3      | QUALITY ASSURANCE, HEALTH CARE/ | 3 MEDLINE | osteoarth*.ti,ab |
| 4      | BENCHMARKING/ | 4 MEDLINE | osteoarth*.ti,ab |
| 5      | CLINICAL AUDIT/ | 5 MEDLINE | osteoarth*.ti,ab |
| 6      | MEDICAL AUDIT/ | 6 MEDLINE | osteoarth*.ti,ab |
| 7      | FACILITY REGULATION AND CONTROL/ | 7 MEDLINE | osteoarth*.ti,ab |
| 8      | GUIDELINES AS TOPIC/ | 8 MEDLINE | osteoarth*.ti,ab |
| 9      | PRACTICE GUIDELINES AS TOPIC/ | 9 MEDLINE | osteoarth*.ti,ab |
| 10     | TOTAL QUALITY MANAGEMENT/ | 10 MEDLINE | osteoarth*.ti,ab |
| 11     | EXP UTILIZATION REVIEW/ | 11 MEDLINE | osteoarth*.ti,ab |
| 12     | EXP "OUTCOME AND PROCESS ASSESSMENT (HEALTH CARE)"/ | 12 MEDLINE | osteoarth*.ti,ab |
| 13     | QUALITY INDICATORS, HEALTH CARE/ | 13 MEDLINE | osteoarth*.ti,ab |
| 14     | 1 OR 2 OR 3 OR 4 OR 5 OR 6 OR 7 OR 8 OR 9 OR 10 OR 11 OR 12 OR 13 | 14 MEDLINE | osteoarth*.ti,ab |
| 15     | osteoarth*.ti,ab | 15 MEDLINE | osteoarth*.ti,ab |
| 16     | osteoarth*.ti,ab | 16 MEDLINE | osteoarth*.ti,ab |
| 17     | 15 OR 16 | 17 MEDLINE | osteoarth*.ti,ab |
| 18     | 14 AND 17 | 18 MEDLINE | osteoarth*.ti,ab |
| 19     | [Limit to: Publication Year 2000-Current and English Language] | 19 MEDLINE | osteoarth*.ti,ab |

The search strategy for use in MEDLINE via NHS Evidence is shown in table 1.

#### Selection of eligible articles

The titles identified were entered in a bibliographical database and duplicates removed. Titles were assessed for relevance by a single reviewer (JJE). The resulting abstracts were evaluated independently by two reviewers (JJE and MK). All those considered relevant by one or both reviewers were entered into the next round. The full texts of the resulting articles were obtained. These were subject to dual independent review of their relevance (JJE plus MK or KSD) and, if there was disagreement on inclusion, by a third reviewer (KSD or MK). This process yielded a final set of articles for the data abstraction round.

#### Method of data extraction

Data extraction forms were designed using the assessment criteria below. The extraction forms were pilot tested and refined by three reviewers. Data were independently extracted by two reviewers (JJE plus MK or KSD). Differences in extraction were resolved by discussion or by a third independent data extraction.

#### Assessment of indicators

The indicators were assessed for quality against criteria used previously, and based on the Outcome Measures in Rheumatology filter (truth, discrimination, feasibility).23–28 The assessment criteria are shown in the online supplementary text S1. Indicators were considered at the level of their development group (for the evidence synthesis, consensus exercise and testing) and at the level of the individual indicator (for discrimination and feasibility).

#### Narrative synthesis

The clinical reviewers (two experienced GPs—JJE and MK—and an academic physiotherapist—KSD) together drafted a narrative synthesis to collate the individual indicators, which was then discussed and revised among all the authors. The indicators were mapped to OA guidance.3–8 Indicator themes developed from the best evidence and consensus method, and rated as feasible for UK primary care, were transformed into a format suitable for implementation. This included a defined numerator (the number of patients receiving a particular element of care) and denominator (those eligible for that element).

### Results

#### Selection of articles

Ten thousand eight hundred and fifty-two unique articles were identified. The final inclusion set numbered 32. There were 10 groups of indicators in 14 development articles, and 18 implementation articles.

The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flow chart setting out the review process can be found in the online supplementary figure S2. Excluded studies are listed in the online supplementary table S3.

#### Assessment of quality

The 10 groups of studies in which indicators for OA care had been developed are listed in table 2. The following aspects of quality assessment were common to all studies and are not included in the table.

Although not every study explicitly declared there to be no conflict of interest, the reviewers considered that no significant resulting bias of the results was likely.

No studies had an identified method of updating the indicators in light of new evidence.

External validity and sensitivity to change had not been demonstrated in any of the indicator development, testing or implementation studies.

Reproducibility, at the level of the individual indicator, is shown in table 3.

Of 10 indicator development study groups, five were based on the Assessing Care of Vulnerable Elders (ACOVE) indicators. Overall, the ACOVE series of indicators were found to most closely fulfil the assessment criteria due to their robust evidence collection and consensus development, and field testing, and update in ACOVE-3. The modifications to ACOVE-1 indicators (for use the English Longitudinal Study of Aging,41 in nursing homes42 and home-based primary care45) were minor, such as to the target population or recommended care process time frames. The degree to which modifications were subject to further empirical study and consensus varied. We judged the modified indicators to be compatible with the originals, although there was variability regarding the indicators of use of oral non-steroidal anti-inflammatories (NSAIDs) and gastroprotective agents, in terms of the drugs recommended or the target population. The RAND indicators were the earliest identified; they were based on a literature review (not identified as systematic) and high-quality consensus exercise. The developers of the Arthritis Foundation indicators had undertaken a ‘comprehensive’ literature review, and a high-quality consensus exercise. One example of implementation (of the non-pharmacological indicators) was found. The remaining indicator sets used an evidence synthesis or consensus exercise which was less rigorous, or not specified. Some had no identified evidence of implementation (eg, Physician Consortium for Performance Improvement (PCPI) indicators).

All identified articles used process-of-care measures as indicators; one indicator set (European Musculoskeletal Conditions Surveillance and Information Network; EUMUSC.net) also used three outcome measures. We identified no papers in which quality improvement over time had been investigated.

---

January 2000 to August 2013. Further studies were identified from other known repositories including the Agency for Healthcare Research and Quality.22

The search strategy for use in MEDLINE via NHS Evidence is shown in table 1.
| Indicator set | Author and date | Truth | Evidence synthesis | Consensus method | Target population | Proposed method of measurement | Testing or implementation | Reliability | Feasibility |
|--------------|----------------|-------|---------------------|------------------|-------------------|--------------------------------|----------------------------|-------------|------------|
| RAND Quality of Care Assessment Tools (RAND QA) | Moore (2000)39 | Literature review, not specified to be systematic | RAND Appropriateness Method | Not specified | Medical record review | Tested in McGlynn et al40 and Asch et al31—in a 4% sample reabstraction reliability was substantial at 3 levels: presence of a condition (k=0.83), indicator eligibility (k=0.76) and indicator scoring (k=0.80) | McGlynn et al40 use a national sample of US citizens in metropolitan areas, using a telephone interview to collect data with subsequent analysis of medical records where consent was given. Asch et al31 use the same data as a comparator for data collected from a random sample of veterans’ health affairs clients and their records |
| ACOVE-1 | MacLean (2001)30 31 | Systematic review supporting indicators produced by a content expert working with a project member knowledgeable about systematic reviews and quality indicator development | Modified RAND/UCLA Appropriateness Method | ‘Vulnerable elders’—persons ≥65 years who are at increased risk for death or functional decline | Not specified | Wenger et al34—10% sample reabstraction: overall error rate was 1.6%; also in Chodosh et al35, Higashi et al36, Ganz et al37, MacLean et al39 and Østerås et al38 used the same data and methodology at a different time point38 Østerås et al40 implemented some of these indicators in a patient self-report format |
| ACOVE-1 adapted for the ELSA | Steel et al (2004)41 | Transposition of previous ACOVE work (referenced). 26 new indicators for the set were suggested by the panel | Modified RAND/UCLA Appropriateness Method | Older patients in the UK ≥65 years) | Interviews for the ELSA | Tested in Steel et al42—(y=0.8, 95% CI=0.7 to 0.9); Østerås et al40 questionnaire-based test-retest k=0.20–0.80, % exact agreement from 62–90% | Steel et al42 and Østerås et al40 separately implemented indicators in UK general practice, using medical record review (computerised and paper notes). Østerås et al40 implemented some of these indicators in a patient self-report format |
| ACOVE-1 adapted for NH implementation (ACOVE/NH) | Saliba et al (2005)42 | Previous referenced (ACOVE) work, plus expert opinion (for modification) and additional indicator development, methodology not specified in detail | Modified Delphi; subsequent overview by ACOVE clinical committee | Long-stay NH residents ≥65 years Exclusions for advanced dementia or poor prognosis | Not specified | Cadogan et al43—tested in 30 nursing homes in California using medical record review. Zingmond et al44 implemented using Medicare and Medicaid eligibility and claims data and a nursing home minimum dataset | Cadogan et al43 implemented indicators in 30 nursing homes in California using medical record review. Zingmond et al44 implemented using Medicare and Medicaid eligibility and claims data and a nursing home minimum dataset |
| ACOVE-1 adapted for the HPCIQ | Smith et al (2007)45 | Based on ACOVE indicators, plus some additional (non-OA) indicators. ACOVE work referenced; additional expert opinion | Modified Delphi techniques | Patients ≥60 years who are homebound | Not specified | No published examples of testing identified | No reliability testing identified |
| ACOVE-3 ACOVE investigators (2007)46 | A systematic review supporting potential indicators produced by a content expert working with | Modified version of the RAND/UCLA Appropriateness Method | Community-dwelling individuals aged ≥65 years who are at greater risk of | Medical records and/ or administrative data, patient or proxy interview | Østerås et al40 | Østerås et al40 questionnaire-based test-retest k=0.20–0.80, % exact agreement from 62–90% | Østerås et al40 implemented some of these indicators in a modified patient self-report format |

Continued
| Indicator set Author and date | Truth Evidence synthesis | Target population | Proposed method of measurement | Testing or implementation | Reliability | Feasibility |
|-------------------------------|--------------------------|-------------------|-------------------------------|--------------------------|-------------|-------------|
| QIGP Underwood (2002)²⁹      | a project member knowledgeable about systematic reviews and indicator development | death or functional decline over a 2-year period | Kirk et al³⁰, Steel et al³², Broadbent et al¹³, Østerås et al⁴⁰ | Tested for non-OA indicators in Kirk et al³⁰—as OA was not included in this exercise, it is not known what degree of reliability exists for these indicators; Østerås et al⁴⁰ questionnaire-based test-retest \( k=0.20-0.80, \% \text{ exact agreement from 62–90\%} \) | Kirk et al³⁰ implemented in 16 UK general practices in two areas using data from electronic and paper records. Steel et al² and Broadbent et al¹³ separately implemented the NSAID indicator in UK general practice, using medical record review (computerised and paper notes). Østerås et al⁴⁰ implemented some of these indicators in a modified patient self-report format |
| Arthritis Foundation Arthritis Foundation 2004₁¹ ²² | Comprehensive literature search and expert opinion | Patients with OA | Li et al³³ | No reliability testing identified | Li et al³³ used a postal survey in Canada (sampling frame from an administrative database in British Columbia) to assess non-pharmacological indicators No feasibility testing identified |
| PCPI (2006)²⁴             | PCPI website refers to a methodology committee but no specific information in the indicator set to identify how it was developed | All patients aged \( \geq 21 \) years with a diagnosis of OA | Medical record data extraction (detailed numerator and denominator information provided) | No published examples of testing identified | No reliability testing identified |
| EUMUSC.net (2012)²⁵       | Developed from the EUMUSC.net standards of care for OA and refined by researchers and patient representatives | All adult patients with OA of hand, hip or knee | Varies. Examples include patient record or survey. Numerator and denominator clearly identified | No published examples of testing identified | No reliability testing identified |

ACOVE, Assessing Care of Vulnerable Elders; DARE, Database of Abstracts of Reviews of Effects; ELSA, English Longitudinal Study of Ageing; EUMUSC.net, European Musculoskeletal Conditions Surveillance and Information Network; HPCQI, Home-based Primary Care Quality Initiative; NH, nursing home; NSAIDs, non-steroidal anti-inflammatories; OA, osteoarthritis; PCPI, Physician Consortium for Performance Improvement; QIGP, Quality Indicators for General Practice; UCLA, University of California, Los Angeles; VE, vulnerable elder.
| Overarching theme (source) | 'Exemplar' indicator | Reproducibility (other sources of similar indicators) | Implementation references and comment on feasibility |
|---------------------------|----------------------|-----------------------------------------------------|-----------------------------------------------------|
| Holistic Assessment: Pain  
(EULAR (all sites), NICE)  | IF a VE has symptomatic OA of the knee or hip, THEN pain should be assessed when new to a primary care or musculoskeletal disease practice and annually… (ACOVE-3)  | RAND QA,29 ACOVE-1,32 33 and as adapted (ELSA,41  
HPCQ(45), Arthritis Foundation,51 52 PCPI,54 EUMUSC.net55  | Requires change in routine coding to improve capture of this information 12 13 30 31 34 35 38 40  |
| Holistic Assessment: Function  
(ACR (hand), EULAR (all sites), NICE )  | IF a VE has symptomatic OA of the knee or hip, THEN functional status should be assessed when new to a primary care or musculoskeletal disease practice and annually… (ACOVE-3)  | RAND QA,29 ACOVE-1,32 33 and as adapted (ELSA,41  
HPCQ(45), Arthritis Foundation,51 52 PCPI,54  | Requires change in routine coding to improve capture of this information 12–14 34 36 40  |
| Education (EULAR (all sites), NICE, OARSI)  | IF a patient has had a diagnosis of symptomatic OA of the knee or hip for >3 months, THEN education about the natural history, treatment, and self-management of OA should have been given or recommended at least once… (Arthritis Foundation)  | ACOVE-1 (2 variations—new and pre-existing disease), 12 33 and as adapted (ELSA41), EUMUSC.net55  | Requires change in routine coding to improve capture of this information  |
| Exercise 1 and 2 (ACR (hip, knee), EULAR (all sites), NICE, OARSI)  | IF an ambulatory VE has symptomatic OA of the knee or hip for longer than 3 months and is able to exercise, THEN a directed or supervised muscle strengthening or aerobic exercise program should be recommended and activity reviewed annually… (ACOVE-3)  | Initial recommendation RAND QA,29 ACOVE-1 (indicators for new and pre-existing disease),12 33 and as adapted (ELSA,41  
ACOVE/NH,42 HPCQ(45), Arthritis Foundation,51 52  
PCPI,54 EUMUSC.net55  | Initial recommendation 14 30 31 34 37 38 40 43 53  
Annual review  |
| Weight loss 1 (ACR (hip, knee), NICE, OARSI)  | IF a VE is obese (body mass index (BMI) ≥30 kg/m²), THEN he or she should be advised annually to lose weight… (ACOVE-3)  | EUMUSC.net55  | Requires change in routine coding to improve capture of this information  |
| Weight loss 2 (ACR (hip, knee), NICE, OARSI)  | IF a patient has symptomatic OA of the knee or hip and is overweight (as defined by body mass index of ≥27 kg/m²), THEN the patient should be advised to lose weight at least annually AND the benefit of weight loss on the symptoms of OA should be explained to the patient… (Arthritis Foundation)  | Arthritis Foundation,51 52  | No implementation studies identified for this indicator. Should be captured from existing weight and health promotion records 40 53  |
| Aids and devices 1 (ACR (hip, knee), EULAR (hip, knee), NICE, OARSI)  | IF a VE has symptomatic OA of the hip or knee and has difficulty walking that makes ADL difficult for longer than 3 months, THEN the need for ambulatory assistive devices should be assessed… (ACOVE-3)  | Arthritis Foundation,51 52 EUMUSC.net55  | Consider a lower BMI threshold of 25 kg/m² for consistency with the usual definition of ‘overweight’. Should be captured from existing weight and health promotion records. 40 53  |
| Aids and devices 2 (ACR (hand), NICE)  | IF a VE has symptomatic OA and has difficulty with non-ambulatory ADL, THEN the need for ADL assistive devices should be assessed… (ACOVE-3)  | Arthritis Foundation,51 52 EUMUSC.net55  | Requires change in routine coding to improve capture of this information 40 53  |
| Paracetamol 1 (ACR (hip, knee), EULAR (all sites), NICE, OARSI)  | IF a VE is started on pharmacological therapy to treat OA, THEN acetaminophen should be tried first… (ACOVE-3)  | RAND QA,29 ACOVE-1,32 33 and as adapted (ELSA,41  
ACOVE/NH,42 HPCQ(45), QiGP,49 Arthritis Foundation,51 52  
PCPI,54 EUMUSC.net55  | Requires change in routine coding to improve capture of this information 12–14 34 36 40 43 50  |
| Paracetamol 2 (ACR (hip, knee), EULAR (all sites), NICE, OARSI)  | IF oral pharmacological therapy for OA is changed from acetaminophen to a different oral agent, THEN there should be evidence that the patient has had a trial of maximum dose acetaminophen (suitable for age/ comorbidities)… (Arthritis Foundation)  | ACOVE-1,32 33 and as adapted (ELSA,41  
ACOVE/NH,42 HPCQ(45),  | Requires change in routine coding to capture over-the-counter drug use 12 13 34 36 43  |
| Oral NSAIDs 1 (all guidance)  | IF NSAIDs are considered, ibuprofen should be considered for first-line treatment unless contraindicated or intolerant.* (QiGP)  | Modifications exist in implementation studies: Steel et al12 Broadbent et al53 to include use of COX-2 selective drugs  | Requires change in routine coding to capture over-the-counter drug use 12 13 50  |
| Oral NSAIDs 2 (all guidance)  | Percentage of patients aged 21 years and older with a diagnosis of OA on prescribed or OTC NSAIDs who were assessed for GI and renal risk factors. (PCPI)  | Two indicators from ACOVE-3 refer to risks from NSAIDs and aspirin to be ‘discussed and documented’; 46–48 EUMUSC.net55  | Requires change in routine coding to capture over-the-counter drug use 12 13  |

Continued
### Table 3

| Overarching theme (source) | Exemplar indicator | Implementation references and comment on feasibility |
|---------------------------|---------------------|-----------------------------------------------------|
| Gastroprotection (EULAR (all sites), NICE, OARSI) | If a VE has severe symptomatic OA of the knee or hip despite non-surgical therapy, THEN a referral to an orthopaedic surgeon should be made. | It would be feasible to capture the presence of non-surgical therapy indicators in the record, though routine data sources may not be used to determine the need for a surgical operation. |
| Specialist referral (EULAR (all sites), NICE, OARSI) | Has any doctor or nurse ever talked to you about: (1) What your arthritis or joint pain is? (2) How your arthritis or joint pain from getting worse?, (3) How your arthritis can be treated? Each of these questions has a yes or no answer. | Evidence from implementation studies suggests that the indicator as worded is less feasible for implementation in primary care, requiring either a more generic indicator or a series of specific patient self-report indicators; we propose a more generic indicator. |
| Exercise and physiotherapy | The ACOVE-3 examples were rated most highly. Exemplar indicators have been selected for these elements of care. Indicators for joint examination and joint aspiration arose less frequently, though were still the result of at least one high quality evidence synthesis and consensus exercise, but had not successfully been implemented. |
| Holistic assessment | There were 28 occurrences of indicators related to holistic assessment of patients. Assessments of pain and function were relatively frequent. The ACOVE-3 examples were rated most highly. Exemplar indicators have been selected for these elements of care. Indicators for joint examination and joint aspiration arose less frequently, though were still the result of at least one high quality evidence synthesis and consensus exercise, but had not successfully been implemented. |

### Narrative synthesis

The indicators identified in the studies were grouped into themes. A summary of exemplar indicators is shown in table 3. The basis of the exemplar choice from the truth and feasibility aspects of the evidence is shown (discrimination was not found to be empirically supported). A list of indicators suitable for routine implementation in primary care is shown in table 4. Online supplementary table S4 lists indicators which do not currently provide sufficient evidence or feasibility for implementation in primary care.

1. Holistic assessment
   - There were 28 occurrences of indicators related to holistic assessment of patients.
   - Assessments of pain and function were relatively frequent. The ACOVE-3 examples were rated most highly. Exemplar indicators have been selected for these elements of care. Indicators for joint examination and joint aspiration arose less frequently, though were still the result of at least one high quality evidence synthesis and consensus exercise, but had not successfully been implemented.

2. Education and information
   - There were 18 occurrences of indicators for education in OA. The Arthritis Foundation indicator was selected due to its cited evidence synthesis and consensus method, and its consistency with the previously implemented ACOVE-1 and recently published EUMUSC.net indicators; no education indicator was included in ACOVE-3. There was some variation in the timeframes for education. It was not clear from most studies implementing this indicator theme how the required level of detail about type of education was obtained. For example, one study asked the patient in a telephone interview “Has any doctor or nurse ever talked to you about: (1) What your arthritis or joint pain will be like as time goes on, or the natural history of arthritis?, (2) How to keep your arthritis or joint pain from getting worse?, (3) How your arthritis can be treated?”; a criterion to pass the indicator was at least one positive response. Evidence from implementation studies suggests that the indicator as worded is less feasible for implementation in primary care, requiring either a more generic indicator or a series of specific patient self-report indicators; we propose a more generic indicator.

3. Exercise and physiotherapy
   - There were 22 occurrences of indicators recommending or prescribing exercise or physiotherapy. One targeted patients with hand, hip and knee OA; one self-report indicator implemented also included patients with hand, hip or knee OA; six refer to exercise for patients with OA of the hip or knee; the remainder specify those with knee OA. There were variations between indicators on exercise, with some recommending that a programme be ‘prescribed’, ‘recommended’ or ‘considered’. Some referred to specific strengthening programmes, others to general aerobic exercise, or physical therapy. For example, one study used a record of prescription for lower extremity strengthening or ambulation with a Physical Therapist or Restorative Nursing Assistant after OA diagnosis as a criterion; others used non-routine sources such as patient interview or unspecified sources. Evidence from implementation studies suggests that feasible indicators for primary care relate to the offer of exercise advice or physiotherapy referral, and review of current exercise activity. It would be feasible to separate two elements of the ACOVE-3 indicator into an indicator for advice,
recommendation or prescription of exercise, and an indicator of annual review of activity.

4. Weight management

There were eight occurrences of indicators regarding weight loss in overweight patients, six for patients with OA and two for primary prevention. There was some variation in the BMI intervention threshold as well as in the type of advice or referral. There were two implementation studies identified, of the Arthritis Foundation indicator regarding weight management in symptomatic OA, in which Li et al.\(^5\) used entry to a weight-loss programme or diététics appointment as criteria for indicator achievement and the weight loss advice self-report indicator in Østerås et al.\(^6\) A primary care indicator related to advice regarding weight loss to reduce the risk of OA, or to improve symptoms in people with established OA would be feasible. A further identified indicator, regarding referral to a weight-loss programme if a person has been overweight for 3 years or more, would be less feasible and desirable, due to greater difficulty establishing the denominator population.

5. Assistive devices (ambulatory and other)

There were nine occurrences of indicators for assessment of need for assistive devices. These covered assessment of need for ambulatory and non-ambulatory assistive devices but there were no specifically recommended interventions. Two examples of implementation were found, of Arthritis Foundation indicators (similar to and consistent with the ACOVE-3 indicators) by Li et al.\(^5\) in which credit was given when a patient had seen a physiotherapist or occupational therapist for ambulatory or non-ambulatory devices respectively within the previous year, and similar patient self-report indicators in Østerås et al.\(^6\) In line with this, general indicators for referral or assessment for ambulatory or assistive devices currently appear feasible in primary care.

6. Analgesics (paracetamol and oral NSAIDs)

There were 53 occurrences of indicators for use of analgesics in OA. These covered topics such as assessment of current use or consideration of analgesics; use of appropriate first-line analgesics; and risk assessment and communication. Preferred indicators generally result from at least one high quality evidence synthesis and consensus exercise, although the basis for the NSAID risk assessment indicator from the PCPI is unclear\(^5\) (though consistent with a similar indicator from the ACOVE-1 group). Where available, the ACOVE-3 indicators were chosen. Several indicators regarding use of paracetamol and NSAIDs are considered feasible for use in primary care (see table 3). Indicators regarding assessment of existing use and consideration of additional treatment from the PCPI\(^5\) are an implemented indicator regarding stronger analgesics (Østerås et al.\(^6\)) were not selected due to an unspecified evidence base and consensus approach; indicators regarding risk explanation were also not selected due to difficulties implementing these in routine data sources (without free text medical record analysis).

7. Gastroprotection

There were 13 occurrences of indicators for use of gastroprotective agents under certain conditions. However, there were variations in the triggers for prescribing a gastroprotective agent, and in the choice of agent to be used. The broadest (PCPI) indicator\(^5\) cites a meta-analysis as having indicated that use of gastrointestinal prophylaxis can be effective in reducing the incidence of adverse events. This would be consistent with the NICE recommendation that everyone over 45 years prescribed a NSAID for OA should be coprescribed a proton pump inhibitor.\(^8\) Where indicators have been implemented, they often use past medical history or co-therapy with other agents (eg, aspirin or warfarin) to determine the denominator group for this indicator. The PCPI indicator is the most feasible, although this has been narrowed to include only proton-pump inhibitor gastroprotection in line with NICE guidance.

8. X-rays, injections, specialist assessment and joint replacement

There were 16 instances of indicators for referral to a specialist and use of X-rays when symptoms were not improving under non-surgical care. As guidance for management of OA does not recommend routine use of X-rays, and no examples of implementation of X-ray indicators was found, this indicator was not considered feasible. A number of indicators referred to failure
of other therapies as a prerequisite for specialist referral but ‘failure’ was not consistently defined. One study asked patients if they had pain and functional impairment, and had been offered a joint replacement or orthopaedic assessment. Another used a patient self-report to identify failure of conservative treatment leading to referral. An indicator mandating that all other indicators must have been recorded as appropriately met prior to referral was considered to be feasible.

There was also one indicator implemented for the consideration of steroid injections for acute symptomatic deterioration. This was not considered feasible for routine implementation in primary care since acute deterioration is hard to identify from the record and many injections take place in secondary care.

9. Outcome measure indicators

The EUMUSC.net project also identified three outcome measures:

- a 20% functional improvement within 3 months of a treatment initiation or change
- a 20% reduction in pain within 3 months of a treatment initiation or change
- enablement of workforce participation for people of working age.

These were considered less feasible for primary care due to the complexity of accounting for comorbidities and case-mix.

**DISCUSSION**

Through a systematic review of OA indicators and a quality appraisal of the indicator development and implementation, we identified 15 indicators of the quality of primary care for OA which could be implemented, benefiting patients, clinicians and policy development.

While the conclusions of the published guidance diverge in some aspects (particularly the use of Symptomatic Slow-Acting Drugs in Osteo-Arthritis, and in some of the detail of oral NSAID use and gastroprotection), the interventions recommended by the different expert groups are broadly similar. The selected indicators were broadly applicable across all the guidance groups.

Within themes, there are differences between some of the identified indicators. Indicators sometimes target differing populations (eg, OA of the knee or any OA), frequency or threshold of assessment or intervention, type of treatment (eg, variation in oral NSAID recommended, and type of gastroprotective agent). These differences are not sufficiently major to cause difficulties in the implementation of the underlying indicator theme.

There are some limitations in this review. There may be indicators not captured by the search strategy (including any prior to 2000, and non-English language indicators). Given the thorough nature of the indicator development methodology for a number of the indicator sets, it seems unlikely that any major themes will have been omitted. In contrast with the assessment of publications on randomised controlled trials (eg, the approach taken by the Cochrane Collaboration), quality assessment of indicators themselves is not a highly developed methodology.

We have selected indicators judged sufficiently robust and feasible for use in routine practice. The use of indicators is dependent upon systematic information capture. In the UK, approximately 90% of prescriptions are obtained with no cost to the patient, and over-the-counter analgesics are restricted in quantity. Analgesics and NSAIDs indicators based on data from computer-generated prescriptions are likely to be valid with no change to recording practice. Other indicators would require a change in coding practice (more detailed coded clinical information). The indicators should be generally applicable to countries with well-developed primary care systems and electronic medical records. The indicators would work best with strategic implementation, for example by inclusion in the Quality and Outcomes Framework of pay-for-performance in UK primary care.

While there are some domains with well-developed and valid indicators, some elements do not have such indicators. For example (1) Holistic assessment: all dimensions other than pain and function, notably periodic review, a jointly formulated management plan and the effect of comorbidities; (2) Education and self-management: the development of a self-management plan and thermotherapy; (3) Non-pharmacological management: manipulation and stretching, electrotherapy, bracing, joint support, footwear and insoles; and (4) Pharmacological management: topical NSAIDs, capsaicin and intra-articular injections. In principle, some of these areas might be suitable for the development of indicators.

We did not identify any negative (‘do not do’) indicators. There are some areas of guidance from which one might usefully derive such indicators for use in primary care. For example, the use of topical rubefacients, electroacupuncture, nutraceuticals, or intra-articular hyaluronan injections, or referral for arthroscopic lavage for OA, based on the NICE guidance.

We found no evidence of external validity (that implementation of indicators is associated with quality improvement). Also, there is no evidence of indicators’ sensitivity to change, so this must currently be assumed. The degree to which a change in recording of the care processes actually reflects a change in the quality of care delivered has not clearly been identified. These areas warrant further investigation: an increased use of patient-reported measures such as those used by Österäs et al would help identify changes in process delivery and outcome.

We have identified a range of indicators for OA which have a good evidence base, are consistent with international guidance, and many of which have been implemented previously. As the disease burden of OA is high, and much of it is presented clinically to GPs, incorporation of these indicators to routine primary care practice is recommended.

**Acknowledgements**

The authors wish to thank the anonymous reviewers of the manuscript for their insightful and helpful observations; also Professor Peter Croft, Professor Danielle van der Windt and Dr Olaakun Utman, of the Arthritis Research UK Primary Care Centre at Keele University for comments on an earlier draft of this manuscript.

**Contributors**

KSD developed the idea for the review. JJE and ILJ designed the search strategy. JJE, MK and KSD analysed the papers and extracted the data. All authors contributed to the analysis and revised the paper. JJE is the guarantor.

**Funding**

This report is independent research arising from an In-Practice Fellowship supported by the National Institute for Health Research (NIHR) with additional funding from the NIHR Programme Grants for Applied Research Programme (Grant Reference Number RP-PG-0407-10386). The views expressed in this publication are those of the authors and not necessarily those of the NHS, the National Institute for Health Research or the Department of Health.

**Competing interests**

All authors have completed the ICMJE uniform disclosure form at http://www.icmje.org/coi_disclosure.pdf (available on request from the corresponding author) and declare: no support from any organisation for the submitted work other than detailed above; JJE, MK and JB provide general medical services and benefit financially from the Quality and Outcomes Framework (which does not currently include osteoarthritis), KSD has been an invited speaker to the EULAR conference and a member of the NICE OA guideline development group— Keele University have received payments and reimbursements of travel and other expenses related to these activities. There are no other financial relationships with any organisations that might have an interest in the submitted work in the previous 3 years. There are no other relationships or activities that could appear to have influenced the submitted work.

Edwards JJ, et al. Ann Rheum Dis 2015;74:490–498. DOI:10.1136/annrheumdis-2013-203913
